



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/12, C07K 14/22, G01N 33/68</b>		<b>A1</b>	(11) International Publication Number: <b>WO 97/32980</b>
			(43) International Publication Date: 12 September 1997 (12.09.97)
(21) International Application Number: PCT/CA97/00163 (22) International Filing Date: 7 March 1997 (07.03.97) (30) Priority Data: 08/613,009      8 March 1996 (08.03.96)      US 08/778,570      3 January 1997 (03.01.97)      US (60) Parent Application or Grant (63) Related by Continuation US      08/778,570 (CIP) Filed on      3 January 1997 (03.01.97) (71) Applicant (for all designated States except US): CONNAUGHT LABORATORIES LIMITED [CA/CA]; 1755 Steeles Avenue West, North York, Ontario M2R 3T4 (CA). (72) Inventors; and (75) Inventors/Applicants (for US only): MYERS, Lisa, E. [CA/CA]; 187 Elizabeth Street, Guelph, Ontario N1E 2X5 (CA). SCHRYVERS, Anthony, B. [CA/CA]; 39 Edforth Road, N.W., Calgary, Alberta T3A 3V8 (CA). HARKNESS, Robin, E. [CA/CA]; Apartment #1706, 640 Sheppard Avenue East, Willowdale, Ontario M2K 1B8 (CA). LOOSMORE, Sheena, M. [CA/CA]; 70 Crawford		Rose Drive, Aurora, Ontario L4G 4R4 (CA). DU, Run-Pan [CA/CA]; 299 Chelwood Drive, Thornhill, Ontario L4J 7Y8 (CA). YANG, Yan-Ping [CA/CA]; Apartment 1709, 120 Torresdale Avenue, Willowdale, Ontario M2R 3N7 (CA). KLEIN, Michel, H. [CA/CA]; 16 Munro Boulevard, Willowdale, Ontario M2P 1B9 (CA). (74) Agent: STEWART, Michael, I.; Sim & McBurney, 6th floor, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, VZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(54) Title: TRANSFERRIN RECEPTOR GENES OF MORAXELLA			
(57) Abstract			
<p>Purified and isolated nucleic acid molecules are provided which encode transferrin receptor proteins of <i>Moraxella</i>, such as <i>M. catarrhalis</i> or a fragment or an analog of the transferrin receptor protein. The nucleic acid sequence may be used to produce recombinant transferrin receptor proteins Tbp1 and Tbp2 of the strain of <i>Moraxella</i> free of other proteins of the <i>Moraxella</i> strain for purposes of diagnostics and medical treatment. Furthermore, the nucleic acid molecule may be used in the diagnosis of infection.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

**TITLE OF INVENTION****TRANSFERRIN RECEPTOR GENES OF MORAXELLA****FIELD OF INVENTION**

5 The present invention relates to the molecular cloning of genes encoding transferrin receptor (TfR) proteins and, in particular, to the cloning of  
10 transferrin receptor genes from *Moraxella* (*Branhamella*) *catarrhalis*.

**REFERENCE TO RELATED APPLICATION**

15 This application is a continuation-in-part of copending United States Patent Application No. 08/778,570 filed January 3, 1997, which itself is a continuation-in-part of United States Patent Application No. 08/613,009 filed March 8, 1996.

**BACKGROUND OF THE INVENTION**

20 *Moraxella* (*Branhamella*) *catarrhalis* bacteria are Gram-negative diplococcal pathogens which are carried asymptotically in the healthy human respiratory tract.

In recent years, *M. catarrhalis* has been recognized as an important causative agent of otitis media. In  
25 addition, *M. catarrhalis* has been associated with sinusitis, conjunctivitis, and urogenital infections, as well as with a number of inflammatory diseases of the lower respiratory tract in children and adults, including pneumonia, chronic bronchitis, tracheitis, and  
30 emphysema (refs. 1 to 8). (Throughout this application, various references are cited in parentheses to describe more fully the state of the art to which this invention pertains. Full bibliographic information for each citation is found at the end of the specification,  
35 immediately preceding the claims. The disclosures of these references are hereby incorporated by reference

**SUBSTITUTE SHEET (RULE 26)**

into the present disclosure). Occasionally, *M. catarrhalis* invades to cause septicaemia, arthritis, endocarditis, and meningitis (refs. 9 to 13).

Otitis media is one of the most common illnesses of early childhood and approximately 80% of all children suffer at least one middle ear infection before the age of three (ref. 14). Chronic otitis media has been associated with auditory and speech impairment in children, and in some cases, has been associated with learning disabilities. Conventional treatments for otitis media include antibiotic administration and surgical procedures, including tonsillectomies, adenoidectomies, and tympanocentesis. In the United States, treatment costs for otitis media are estimated to be between one and two billion dollars per year.

In otitis media cases, *M. catarrhalis* commonly is co-isolated from middle ear fluid along with *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*, which are believed to be responsible for 50% and 30% of otitis media infections, respectively. *M. catarrhalis* is believed to be responsible for approximately 20% of otitis media infections (ref. 15).

Epidemiological reports indicate that the number of cases of otitis media attributable to *M. catarrhalis* is increasing, along with the number of antibiotic-resistant isolates of *M. catarrhalis*. Thus, prior to 1970, no  $\beta$ -lactamase-producing *M. catarrhalis* isolates had been reported, but since the mid-seventies, an increasing number of  $\beta$ -lactamase-expressing isolates have been detected. Recent surveys suggest that 75% of clinical isolates produce  $\beta$ -lactamase (ref. 16, 26).

Iron is an essential nutrient for the growth of many bacteria. Several bacterial species, including *M. catarrhalis*, obtain iron from the host by using transferrin receptor proteins to capture transferrin. A number of bacteria including *Neisseria meningitidis*

(ref. 17), *N. gonorrhoeae* (ref. 18), *Haemophilus influenzae* (ref. 19), as well as *M. catarrhalis* (ref. 20), produce outer membrane proteins which specifically bind human transferrin. The expression of these proteins is regulated by the amount of iron in the environment.

The two transferrin receptor proteins of *M. catarrhalis*, designated transferrin binding protein 1 (Tbp1) and transferrin binding protein 2 (Tbp2), have molecular weights of 115 kDa (Tbp1) and approximately 80 to 90 kDa (Tbp2). Unlike the transferrin receptor proteins of other bacteria which have an affinity for apotransferrin, the *M. catarrhalis* Tbp2 receptors have a preferred affinity for iron-saturated (i.e., ferri-) transferrin (ref. 21).

*M. catarrhalis* infection may lead to serious disease. It would be advantageous to provide a recombinant source of transferrin binding proteins as antigens in immunogenic preparations including vaccines, carriers for other antigens and immunogens and the generation of diagnostic reagents. The genes encoding transferrin binding proteins and fragments thereof are particularly desirable and useful in the specific identification and diagnosis of *Moraxella* and for immunization against disease caused by *M. catarrhalis* and for the generation of diagnostic reagents.

#### SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules encoding a transferrin receptor of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein. The nucleic acid molecules provided herein are useful for the specific detection of strains of *Moraxella* and for diagnosis of infection by *Moraxella*. The purified and isolated nucleic acid

molecules provided herein, such as DNA, are also useful for expressing the *tbp* genes by recombinant DNA means for providing, in an economical manner, purified and isolated transferrin receptor proteins as well as subunits, fragments or analogs thereof. The transferrin receptor, subunits or fragments thereof or analogs thereof, as well as nucleic acid molecules encoding the same and vectors containing such nucleic acid molecules, are useful in immunogenic compositions for vaccinating against diseases caused by *Moraxella*, the diagnosis of infection by *Moraxella* and as tools for the generation of immunological reagents. Monoclonal antibodies or mono-specific antisera (antibodies) raised against the transferrin receptor protein, produced in accordance with aspects of the present invention, are useful for the diagnosis of infection by *Moraxella*, the specific detection of *Moraxella* (in, for example, *in vitro* and *in vivo* assays) and for the treatment of diseases caused by *Moraxella*.

In accordance with one aspect of the present invention, there is provided a purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella*, more particularly, a strain of *M. catarrhalis*, specifically *M. catarrhalis* strain 4223, Q8 or R1, or a fragment or an analog of the transferrin receptor protein.

In one preferred embodiment of the invention, the nucleic acid molecule may encode only the Tbp1 protein of the *Moraxella* strain or only the Tbp2 protein of the *Moraxella* strain. In another preferred embodiment of the invention, the nucleic acid may encode a fragment of the transferrin receptor protein of a strain of *Moraxella* having an amino acid sequence which is conserved.

In another aspect of the present invention, there is provided a purified and isolated nucleic acid

molecule having a DNA sequence selected from the group consisting of (a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto; (b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and (c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b). The DNA sequence defined in (c) preferably has at least about 90% sequence identity with any one of the DNA sequences defined in (a) and (b). The DNA sequence defined in (c) may be that encoding the equivalent transferrin receptor protein from another strain of *Moraxella*.

In an additional aspect, the present invention includes a vector adapted for transformation of a host, comprising a nucleic acid molecule as provided herein and may have the characteristics of a nucleotide sequence contained within vectors LEM3-24, pLEM3, pLEM25, pLEM23, SLRD-A, DS-1698-1-1, DS-1754-1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5.

The vector may be adapted for expression of the encoded transferrin receptor, fragments or analogs thereof, in a heterologous or homologous host, in either a lipidated or non-lipidated form. Accordingly, a further aspect of the present invention provides an expression vector adapted for transformation of a host comprising a nucleic acid molecule as provided herein and expression means operatively coupled to the nucleic acid molecule for expression by the host of the transferrin receptor protein or the fragment or analog of the transferrin receptor protein. In specific embodiments of this aspect of the invention, the nucleic acid molecule may encode substantially all the transferrin receptor protein, only the Tbp1 protein,

only the Tbp2 protein of the *Moraxella* strain or fragments of the Tbp1 or Tbp2 proteins. The expression means may include a promoter and a nucleic acid portion encoding a leader sequence for secretion from the host of the transferrin receptor protein or the fragment or the analog of the transferrin receptor protein. The expression means also may include a nucleic acid portion encoding a lipitation signal for expression from the host of a lipidated form of the transferrin receptor protein or the fragment or the analog of the transferrin receptor protein. The host may be selected from, for example, *Escherichia coli*, *Bordetella*, *Bacillus*, *Haemophilus*, *Moraxella*, fungi, yeast or baculovirus and Semliki Forest virus expression systems may be used. In a particular embodiment, the plasmid adapted for expression of Tbp1 is pLEM29 and that for expression of Tbp2 is pLEM33. Further vectors include pLEM-37, SLRD35-A and SLRD-35-B.

In an additional aspect of the invention, there is provided a transformed host containing an expression vector as provided herein. The invention further includes a recombinant transferrin receptor protein or fragment or analog thereof of a strain of *Moraxella* producible by the transformed host.

Such recombinant transferrin receptor protein may be provided in substantially pure form according to a further aspect of the invention, which provides a method of forming a substantially pure recombinant transferrin receptor protein, which comprises growing the transformed host provided herein to express a transferrin receptor protein as inclusion bodies, purifying the inclusion bodies free from cellular material and soluble proteins, solubilizing transferrin receptor protein from the purified inclusion bodies, and purifying the transferrin receptor protein free from other solubilized materials. The substantially pure



recombinant transferrin receptor protein may comprise Tbp1 alone, Tbp2 alone or a mixture thereof. The recombinant protein is generally at least about 70% pure, preferably at least about 90% pure.

5 Further aspects of the present invention, therefore, provide recombinantly-produced Tbp1 protein of a strain of *Moraxella* devoid of the Tbp2 protein of the *Moraxella* strain and any other protein of the *Moraxella* strain and recombinantly-produced Tbp2 protein  
10 of a strain of *Moraxella* devoid of the Tbp1 protein of the *Moraxella* strain and any other protein of the *Moraxella* strain. The *Moraxella* strain may be *M. catarrhalis* 4223 strain, *M. catarrhalis* Q8 strain or *M. catarrhalis* R1 strain.

15 In accordance with another aspect of the invention, an immunogenic composition is provided which comprises at least one active component selected from at least one nucleic acid molecule as provided herein and at least one recombinant protein as provided herein, and a  
20 pharmaceutically acceptable carrier therefor or vector therefor. The at least one active component produces an immune response when administered to a host.

The immunogenic compositions provided herein may be formulated as vaccines for *in vivo* administration to a  
25 host. For such purpose, the compositions may be formulated as a microparticle, capsule, ISCOM or liposome preparation. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to  
30 mucosal surfaces. The immunogenic compositions of the invention (including vaccines) may further comprise at least one other immunogenic or immunostimulating material and the immunostimulating material may be at least one adjuvant or at least one cytokine. Suitable  
35 adjuvants for use in the present invention include (but are not limited to) aluminum phosphate, aluminum

hydroxide, QS21, Quil A, derivatives and components thereof, ISCOM matrix, calcium phosphate, calcium hydroxide, zinc hydroxide, a glycolipid analog, an octadecyl ester of an amino acid, a muramyl dipeptide polyphosphazene, ISCOPREP, DC-chol, DDBA and a lipoprotein. Advantageous combinations of adjuvants are described in copending United States Patent Applications Nos. 08/261,194 filed June 16, 1994 and 08/483,856, filed June 7, 1995, assigned to the assignee hereof and the disclosures of which are incorporated herein by reference thereto (WO 95/34308).

In accordance with another aspect of the invention, there is provided a method for generating an immune response in a host, comprising the step of administering to a susceptible host, such as a human, an effective amount of the immunogenic composition provided herein. The immune response may be a humoral or a cell-mediated immune response and may provide protection against disease caused by *Moraxella*. Hosts in which protection against disease may be conferred include primates, including humans.

In a further aspect, there is provided a live vector for delivery of transferrin receptor to a host, comprising a vector containing the nucleic acid molecule as described above. The vector may be selected from *Salmonella*, BCG, adenovirus, poxvirus, vaccinia and poliovirus.

The nucleic acid molecules provided herein are useful in diagnostic applications. Accordingly, in a further aspect of the invention, there is provided a method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with a nucleic acid molecule as provided herein to produce duplexes comprising the nucleic acid molecule and any nucleic

acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining the production of the duplexes.

5 In addition, the present invention provides a diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

(a) a nucleic acid molecule as provided herein;

10 (b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any such nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

15 (c) means for determining production of the duplexes.

The invention further includes the use of the nucleic acid molecules and proteins provided herein as medicines. The invention additionally includes the use  
20 of the nucleic acid molecules and proteins provided herein in the manufacture of medicaments for protection against infection by strains of *Moraxella*.

Advantages of the present invention include:

- an isolated and purified nucleic acid  
25 molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein;

- recombinantly-produced transferrin receptor proteins, including Tbp1 and Tbp2, free from each other  
30 and other *Moraxella* proteins; and

- diagnostic kits and immunological reagents for specific identification of *Moraxella*.

#### BRIEF DESCRIPTION OF DRAWINGS

35 The present invention will be further understood from the following description with reference to the

drawings, in which:

Figure 1 shows the amino acid sequences (SEQ ID Nos: 17 and 18) of a conserved portion of Tbp1 proteins used for synthesis of degenerate primers used for PCR amplification of a portion of the *M. catarrhalis* 4223 *tbpA* gene;

Figure 2 shows a restriction map of clone LEM3-24 containing the *tbpA* and *tbpB* genes from *M. catarrhalis* isolate 4223;

Figure 3 shows a restriction map of the *tbpA* gene for *M. catarrhalis* 4223;

Figure 4 shows a restriction map of the *tbpB* gene for *M. catarrhalis* 4223;

Figure 5 shows the nucleotide sequence of the *tbpA* gene (SEQ ID No: 1 - entire sequence and SEQ ID No: 2 - coding sequence) and the deduced amino acid sequence of the Tbp1 protein from *M. catarrhalis* 4223 (SEQ ID No: 9 - full length and SEQ ID No: 10 - mature protein). The leader sequence (SEQ ID No: 19) is shown by underlining;

Figure 6 shows the nucleotide sequence of the *tbpB* gene (SEQ ID No: 3 - entire sequence and SEQ ID No: 4 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein from *M. catarrhalis* 4223 (SEQ ID Nos: 11 - full length and SEQ ID No: 12 - mature protein). The leader sequence (SEQ ID No: 20) is shown by underlining;

Figure 7 shows a restriction map of clone SLRD-A containing the *tbpA* and *tbpB* genes from *M. catarrhalis* Q8;

Figure 8 shows a restriction map of the *tbpA* gene from *M. catarrhalis* Q8;

Figure 9 shows a restriction map of the *tbpB* gene from *M. catarrhalis* Q8;

Figure 10 shows the nucleotide sequence of the *tbpA* gene (SEQ. ID No: 5 - entire sequence and SEQ ID No: 6 - coding sequence) and the deduced amino acid sequence of

the Tbp1 protein from *M. catarrhalis* Q8 (SEQ ID No: 13 - full length and SEQ ID No: 14 - mature protein);

Figure 11 shows the nucleotide sequence of the *tbpB* gene (SEQ. ID No: 7 - entire sequence and SEQ ID No: 8 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein from *M. catarrhalis* Q8 (SEQ ID No: 15 - full length and SEQ ID No: 16 - mature protein);

Figure 12 shows a comparison of the amino acid sequences of Tbp1 from *M. catarrhalis* strain 4223 (SEQ ID No: 9) and Q8 (SEQ ID No: 13), *H. influenzae* strain Eagan (SEQ ID No: 21), *N. meningitidis* strains B16B6 (SEQ ID No: 22) and M982 (SEQ ID No: 23), and *N. gonorrhoeae* strain FA19 (SEQ ID No: 24). Dots indicate identical residues and dashes have been inserted for maximum alignment;

Figure 13 shows a comparison of the amino acid sequences of Tbp2 from *M. catarrhalis* isolate 4223 (SEQ ID No: 11) and Q8 (SEQ ID No: 15), *H. influenzae* strain Eagan (SEQ ID No: 25), *N. meningitidis* strains B16B6 (SEQ ID No: 26) and M918 (SEQ ID No: 27), and *N. gonorrhoeae* strain FA19 (SEQ ID No: 28). Dots indicate identical residues and dashes have been inserted for maximum alignment;

Figure 14 shows the construction of plasmid pLEM29 for expression of recombinant Tbp1 protein from *E. coli*;

Figure 15 shows an SDS-PAGE analysis of the expression of Tbp1 protein by *E. coli* cells transformed with plasmid pLEM29;

Figure 16 shows a flow chart for purification of recombinant Tbp1 protein;

Figure 17 shows an SDS-PAGE analysis of purified recombinant Tbp1 protein;

Figure 18 shows the construction of plasmid pLEM33 and pLEM37 for expression of TbpA gene from *M. catarrhalis* 4223 in *E. coli* without and with a leader sequence respectively;

Figure 19 shows an SDS-PAGE analysis of the expression of rTbp2 protein by *E. coli* cells transformed with plasmid pLEM37;

Figure 20 shows the construction of plasmid sLRD35B for expression of the *tbpB* gene from *M. catarrhalis* Q8 in *E. coli* without a leader sequence, and the construction of plasmid SLRD35A for expression of the *tbpB* gene from *M. catarrhalis* Q8 in *E. coli* with a leader sequence. Restriction site B = BamHI; Bg = Bgl II; H = Hind III; R = EcoRI;

Figure 21 shows SDS PAGE analysis of the expression of rTbp2 protein in *E. coli* cells, transformed with plasmids SLRD35A and SLRD35B;

Figure 22 shows a flow chart for purification of recombinant Tbp2 protein from *E. coli*;

Figure 23, which includes Panels A and B, shows an SDS-PAGE analysis of the purification of recombinant Tbp2 protein from *M. catarrhalis* strains 4223 (Panel A) and Q8 (Panel B) from expression in *E. coli*;

Figure 24 shows the binding of Tbp2 to human transferrin;

Figure 25, which includes Panels A, B and C, shows the antigenic conservation of Tbp2 protein amongst strains of *M. catarrhalis*;

Figure 26 shows a restriction map of the *tbpB* gene for *M. catarrhalis* R1;

Figure 27 shows the nucleotide sequence of the *tbpB* gene (SEQ ID No: 45 - entire sequence and SEQ ID No: 46 - coding sequence) and the deduced amino acid sequence of the Tbp2 protein of *M. catarrhalis* R1 (SEQ ID No: 47); and

Figure 28 shows a comparison of the amino acid sequences of Tbp2 for *M. catarrhalis* 4223 (SEQ ID No: 21), Q8 (SEQ ID No: 15) and R1 (SEQ ID No: 47). Dots indicate identical residues and dashes have been inserted for maximum alignment. The asterisks indicate

stop codons.

#### GENERAL DESCRIPTION OF THE INVENTION

Any *Moraxella* strain may be conveniently used to provide the purified and isolated nucleic acid, which may be in the form of DNA molecules, comprising at least a portion of the nucleic acid coding for a transferrin receptor as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection.

In this application, the terms "transferrin receptor" (TfR) and "transferrin binding proteins" (Tbp) are used to define a family of Tbp1 and/or Tbp2 proteins which includes those having variations in their amino acid sequences including those naturally occurring in various strains of, for example, *Moraxella*. The purified and isolated DNA molecules comprising at least a portion coding for transferrin receptor of the present invention also include those encoding functional analogs of transferrin receptor proteins Tbp1 and Tbp2 of *Moraxella*. In this application, a first protein is a "functional analog" of a second protein if the first protein is immunologically related to and/or has the same function as the second protein. The functional analog may be, for example, a fragment of the protein, or a substitution, addition or deletion mutant thereof.

Chromosomal DNA from *M. catarrhalis* 4223 was digested with *Sau3A* in order to generate fragments within a 15 to 23 kb size range, and cloned into the *Bam*HI site of the lambda vector EMBL3. The library was screened with anti-Tbp1 guinea pig antisera, and a positive clone LEM3-24, containing an insert approximately 13.2 kb in size was selected for further analysis. Lysate from *E. coli* LE392 infected with LEM3-24 was found to contain a protein approximately 115 kDa

in size, which reacted on Western blots with anti-Tbp1 antisera. A second protein, approximately 80 kDa in size, reacted with the anti-Tbp2 guinea pig antisera on Western blots.

5 In order to localize the *tbpA* gene on the 13.2 kb insert of LEM3-24, degenerate PCR primers were used to amplify a small region of the putative *tbpA* gene of *M. catarrhalis* 4223. The sequences of the degenerate oligonucleotide primers were based upon conserved amino  
10 acid sequences within the Tbp1 proteins of several *Neisseria* and *Haemophilus* species and are shown in Figure 1 (SEQ ID Nos: 17 and 18). A 300 base-pair amplified product was generated and its location within the 4223 *tbpA* gene is indicated by bold letters in  
15 Figure 5 (SEQ ID No: 29). The amplified product was subcloned into the vector pCRII, labelled, and used to probe a Southern blot containing restriction-endonuclease digested clone LEM3-24 DNA. The probe hybridized to a 3.8 kb *HindIII*-*HindIII*, a 2.0 kb *AvrII*-*AvrII*, and 4.2 kb *SalI*-*SphI* fragments (Figure 2).

20 The 3.8 kb *HindIII*-*HindIII* fragment was subcloned into pACYC177, and sequenced. A large open reading frame was identified, and subsequently found to contain approximately 2 kb of the putative *tbpA* gene. The  
25 remaining 1 kb of the *tbpA* gene was obtained by subcloning an adjacent downstream *HindIII*-*HindIII* fragment into vector pACYC177. The nucleotide sequence of the *tbpA* gene from *M. catarrhalis* 4223 (SEQ ID Nos: 1 and 2), and the deduced amino acid sequence (SEQ ID  
30 No: 9 - full length; SEQ ID No: 10 mature protein ) are shown in Figure 5.

35 Chromosomal DNA from *M. catarrhalis* strain Q8 was digested with *Sau3A* I and 15-23 kb fragments were ligated with *BamHI* arms of EMBL3. A high titre library was generated in *E. coli* LE392 cells and was screened using oligonucleotide probes based on the 4223 *tbpA*



sequence. Phage DNA was prepared and restriction enzyme analysis revealed that inserts of about 13-15 kb had been cloned. Phage clone SLRD-A was used to subclone fragments for sequence analysis. A cloning vector (pSKMA) was generated to facilitate cloning of the fragments and plasmids pSLRD1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5 were generated which contain all of *tbpA* and most of *tbpB*. The nucleotide (SEQ ID Nos: 5 and 6) and deduced amino acid sequence (SEQ ID No: 13 - full length, SEQ ID No: 14 - mature protein) of the *tbpA* gene from strain Q8 are shown in Figure 10.

The deduced amino acid sequences for the Tbp1 protein encoded by the *tbpA* genes were found to share some homology with the amino acid sequences encoded by genes from a number of *Neisseria* and *Haemophilus* species (Figure 12; SEQ ID Nos: 21, 22, 23 and 24).

Prior to the present discovery, *tbpA* genes identified in species of *Neisseria*, *Haemophilus*, and *Actinobacillus* have been found to be preceded by a *tbpB* gene with several conserved regions. The two genes typically are separated by a short intergenic sequence. However, a *tbpB* gene was not found upstream of the *tbpA* gene in *M. catarrhalis* 4223. In order to localize the *tbpB* gene within the 13.2 kb insert of clone LEM3-24, a denenerate oligonucleotide probe was synthesized based upon an amino acid sequence EGGFYGP (SEQ ID No: 30), conserved among Tbp2 proteins of several species. The oligonucleotide was labelled and used to probe a Southern blot containing different restriction endonuclease fragments of clone LEM3-24. The probe hybridized to a 5.5 kb *NheI*-*SalI* fragment, which subsequently was subcloned into pBR328, and sequenced. The fragment contained most of the putative *tbpB* gene, with the exception of the promoter region. The clone LEM3-24 was sequenced to obtain the remaining upstream sequence. The *tbpB* gene was located approximately 3 kb

downstream from the end of the *tbpA* gene, in contrast to the genetic organization of the *tbpA* and *tbpB* genes in *Haemophilus* and *Neisseria*. The nucleotide sequence (SEQ ID Nos: 3 and 4) of the *tbpB* gene from *M. catarrhalis* 4223 and the deduced amino acid sequence (SEQ ID Nos: 11, 12) are shown in Figure 6. The *tbpB* gene from *M. catarrhalis* Q8 was also cloned and sequenced. The nucleotide sequence (SEQ ID Nos: 7 and 8) and the deduced amino acid sequence (SEQ ID Nos: 15 and 16) are shown in Figure 11. The *tbpB* gene from *M. catarrhalis* R1 was also cloned and sequenced. The nucleotide sequence (SEQ ID Nos: 45 and 46) and the deduced amino acid sequence (SEQ ID No: 47) are shown in Figure 27. Regions of homology are evident between the *M. catarrhalis* Tbp2 amino acid sequences as shown in the comparative alignment of Figure 28 (SEQ ID Nos: 11, 15 and 47) and between the *M. catarrhalis* Tbp2 amino acid sequences and the Tbp2 sequences of a number of *Neisseria* and *Haemophilus* species, as shown in the comparative alignment in Figure 13 (SEQ ID Nos: 25, 26, 27, 28).

Cloned *tbpA* and *tbpB* genes were expressed in *E. coli* to produce recombinant Tbp1 and Tbp2 proteins free of other *Moraxella* proteins. These recombinant proteins were purified and used for immunization.

The antigenic conservation of Tbp2 protein amongst strains of *M. catarrhalis* was demonstrated by separation of the proteins in whole cell lysates of *M. catarrhalis* or strains of *E. coli* expressing recombinant Tbp2 proteins by SDS PAGE and antiserum immunoblotting with anti-4223 rTbp2 antiserum or anti-Q8 rTbp2 antiserum raised in guinea pigs. *M. catarrhalis* strains 3, 56, 135, 585, 4223, 5191, 8185 and ATCC 25240 were tested in this way and all showed specific reactivity with anti-4223 rTbp2 or anti-Q8 rTbp2 antibody (Figure 25).

In addition, the ability of anti-rTbp2 antibodies from one strain to recognize native or recombinant protein from the homologous or heterologous strain by ELISA is shown in Table 1 below.

5 Amino acid sequencing of the N-termini and cyanogen bromide fragments of transferrin receptor from *M. catarrhalis* 4223 was undertaken. Both N-termini of Tbp1 and Tbp2 were blocked. The putative signal sequences of Tbp1 and Tbp2 are indicated by underlining in Figures 5 and 6 (SEQ ID Nos: 19 and 20) respectively. The deduced amino acid sequences for the N-terminal region of Tbp2 suggests a lipoprotein structure.

10 Results shown in Tables 1 and 2 below illustrate the ability of anti-Tbp1 and anti-Tbp2 guinea pig antisera, produced by the immunization with Tbp1 or Tbp2, to lyse *M. catarrhalis*. The results show that the antisera produced by immunization with Tbp1 or Tbp2 protein isolated from *M. catarrhalis* isolate 4223 were bactericidal against a homologous non-clumping *M. catarrhalis* strain RH408 (a strain previously deposited in connection with United States Patent Application No. 08/328,589, assigned to the assignee hereof, (WO 96/12733) with the American Type Culture Collection, located at 1301 Parklawn Drive, Rockville, Maryland 20852, USA under the terms of the Budapest Treaty on December 13, 1994 under ATCC Deposit No. 55,637) derived from isolate 4223. In addition, antisera produced by immunization with Tbp1 protein isolated from *M. catarrhalis* 4223 were bactericidal against the heterologous non-clumping strain Q8 (a gift from Dr. M.G. Bergeron, Centre Hospitalier de l'Université Laval, St. Foy, Quebec). In addition, antiserum raised against recombinant Tbp2 (rTbp2) protein was bacteriacidal against the homologous strain of *M. catarrhalis*.

35 The ability of isolated and purified transferrin binding proteins to generate bactericidal antibodies is

in vivo evidence of utility of these proteins as vaccines to protect against disease caused by *Moraxella*.

Thus, in accordance with another aspect of the present invention, there is provided a vaccine against infection caused by *Moraxella* strains, comprising an immunogenically-effective amount of a transferrin binding protein from a strain of *Moraxella* and a physiologically-acceptable carrier therefor. Vaccine preparations may comprise antigenically or sequence divergent transferrin binding proteins.

The transferrin binding protein provided herein is useful as a diagnostic reagent, as an antigen for the generation of anti-transferrin protein binding antibodies, as an antigen for vaccination against the disease caused by species of *Moraxella* and for detecting infection by *Moraxella* and other such bacteria.

The transferrin binding protein provided herein may also be used as a carrier protein for haptens, polysaccharides or peptides to make conjugate vaccines against antigenic determinants unrelated to transferrin binding proteins. In additional embodiments of the present invention, therefore, the transferrin binding protein as provided herein may be used as a carrier molecule to prepare chimeric molecules and conjugate vaccines (including glycoconjugates) against pathogenic bacteria, including encapsulated bacteria. Thus, for example, glycoconjugates of the present invention may be used to confer protection against disease and infection caused by any bacteria having polysaccharide antigens including lipooligosaccharides (LOS) and PRP. Such bacterial pathogens may include, for example, *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Escherichia coli*, *Neisseria meningitidis*, *Salmonella typhi*, *Streptococcus mutans*, *Cryptococcus neoformans*, *Klebsiella*, *Staphylococcus aureus* and *Pseudomonas aeruginosa*. Particular antigens which can be conjugated

to transferrin binding protein and methods to achieve such conjugations are described in U.S. Patent Application No. 08/433,522 filed November 23, 1993 (WO 94/12641), assigned to the assignee hereof and the disclosure of which is hereby incorporated by reference thereto.

In another embodiment, the carrier function of transferrin binding protein may be used, for example, to induce an immune response against abnormal polysaccharides of tumour cells, or to produce anti-tumour antibodies that can be conjugated to chemotherapeutic or bioactive agents.

The invention extends to transferrin binding proteins from *Moraxella catarrhalis* for use as an active ingredient in a vaccine against disease caused by infection with *Moraxella*. The invention also extends to a pharmaceutical vaccinal composition containing transferrin binding proteins from *Moraxella catarrhalis* and optionally, a pharmaceutically acceptable carrier and/or diluent.

In a further aspect the invention provides the use of transferrin binding proteins for the preparation of a pharmaceutical vaccinal composition for immunization against disease caused by infection with *Moraxella*.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, treatment of, for example, *Moraxella* infections and the generation of immunological and other diagnostic reagents. A further non-limiting discussion of such uses is further presented below.

#### 1. Vaccine Preparation and Use

Immunogenic compositions, suitable to be used as vaccines, may be prepared from immunogenic transferrin receptor proteins, analogs and fragments thereof encoded by the nucleic acid molecules as well as the nucleic

acid molecules disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-transferrin receptor antibodies and antibodies that are opsonizing or bactericidal. Should the vaccinated  
5 subject be challenged by *Moraxella*, the antibodies bind to the transferrin receptor and thereby prevent access of the bacteria to an iron source which is required for viability. Furthermore, opsonizing or bactericidal anti-transferrin receptor antibodies may also provide  
10 protection by alternative mechanisms.

Immunogenic compositions, including vaccines, may be prepared as injectables, as liquid solutions or emulsions. The transferrin receptor proteins, analogs and fragments thereof and encoding nucleic acid  
15 molecules may be mixed with pharmaceutically acceptable excipients which are compatible with the transferrin receptor proteins, fragments, analogs or nucleic acid molecules. Such excipients may include water, saline, dextrose, glycerol, ethanol, and combinations thereof.

20 The immunogenic compositions and vaccines may further contain auxiliary substances, such as wetting or emulsifying agents, pH buffering agents, or adjuvants, to enhance the effectiveness of the vaccines.

Immunogenic compositions and vaccines may be  
25 administered parenterally, by injection subcutaneously, intradermally or intramuscularly. Alternatively, the immunogenic compositions provided according to the present invention, may be formulated and delivered in a manner to evoke an immune response at mucosal surfaces.

30 Thus, the immunogenic composition may be administered to mucosal surfaces by, for example, the nasal or oral (intragastic) routes. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to  
35 mucosal surfaces. Some such targeting molecules include vitamin B12 and fragments of bacterial toxins, as

described in WO 92/17167 (Biotech Australia Pty. Ltd.), and monoclonal antibodies, as described in U.S. Patent No. 5,194,254 (Barber et al). Alternatively, other modes of administration, including suppositories and oral formulations, may be desirable. For suppositories, binders and carriers may include, for example, polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example, pharmaceutical grades of saccharine, cellulose and magnesium carbonate. These compositions may take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 1 to 95% of the transferrin receptor proteins, fragments, analogs and/or nucleic acid molecules.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective, protective and immunogenic. The quantity to be administered depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and, if needed, to produce a cell-mediated immune response. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the transferrin receptor proteins, analogs and fragments thereof and/or nucleic acid molecules. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and will vary according to the size of the host.

The nucleic acid molecules encoding the transferrin

receptor of *Moraxella* may be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus containing the nucleic acid molecules. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system is contained in, for example, O'Hagan (ref 22). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulmer et al. (ref. 23).

Immunogenicity can be significantly improved if the antigens are co-administered with adjuvants, commonly used as an 0.05 to 1.0 percent solution in phosphate - buffered saline. Adjuvants enhance the immunogenicity of an antigen but are not necessarily immunogenic themselves. Adjuvants may act by retaining the antigen locally near the site of administration to produce a depot effect facilitating a slow, sustained release of antigen to cells of the immune system. Adjuvants can also attract cells of the immune system to an antigen depot and stimulate such cells to elicit immune responses.

Immunostimulatory agents or adjuvants have been used for many years to improve the host immune responses to, for example, vaccines. Intrinsic adjuvants, such as lipopolysaccharides, normally are the components of killed or attenuated bacteria used as vaccines. Extrinsic adjuvants are immunomodulators which are typically non-covalently linked to antigens and are formulated to enhance the host immune responses. Thus, adjuvants have been identified that enhance the immune response to antigens delivered parenterally. Some of these adjuvants are toxic, however, and can cause undesirable side-effects, making them unsuitable for use



in humans and many animals. Indeed, only aluminum hydroxide and aluminum phosphate (collectively commonly referred to as alum) are routinely used as adjuvants in human and veterinary vaccines. The efficacy of alum in increasing antibody responses to diphtheria and tetanus toxoids is well established and an HBsAg vaccine has been adjuvanted with alum. While the usefulness of alum is well established for some applications, it has limitations. For example, alum is ineffective for influenza vaccination and inconsistently elicits a cell mediated immune response. The antibodies elicited by alum-adjuvanted antigens are mainly of the IgG1 isotype in the mouse, which may not be optimal for protection by some vaccinal agents.

A wide range of extrinsic adjuvants can provoke potent immune responses to antigens. These include saponins complexed to membrane protein antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria and mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, and liposomes.

To efficiently induce humoral immune responses (HIR) and cell-mediated immunity (CMI), immunogens are often emulsified in adjuvants. Many adjuvants are toxic, inducing granulomas, acute and chronic inflammations (Freund's complete adjuvant, FCA), cytotoxicity (saponins and pluronic polymers) and pyrogenicity, arthritis and anterior uveitis (LPS and MDP). Although FCA is an excellent adjuvant and widely used in research, it is not licensed for use in human or veterinary vaccines because of its toxicity.

Desirable characteristics of ideal adjuvants include:

- (1) lack of toxicity;
- (2) ability to stimulate a long-lasting immune

response;

(3) simplicity of manufacture and stability in long-term storage;

(4) ability to elicit both CMI and HIR to antigens administered by various routes, if required;

(5) synergy with other adjuvants;

(6) capability of selectively interacting with populations of antigen presenting cells (APC);

(7) ability to specifically elicit appropriate  $T_H1$  or  $T_H2$  cell-specific immune responses; and

(8) ability to selectively increase appropriate antibody isotype levels (for example, IgA) against antigens.

U.S. Patent No. 4,855,283 granted to Lockhoff et al on August 8, 1989, which is incorporated herein by reference thereto, teaches glycolipid analogues including N-glycosylamides, N-glycosylureas and N-glycosylcarbamates, each of which is substituted in the sugar residue by an amino acid, as immuno-modulators or adjuvants. Thus, Lockhoff et al. 1991 (ref. 24) reported that N-glycolipid analogs displaying structural similarities to the naturally-occurring glycolipids, such as glycopospholipids and glyco glycerolipids, are capable of eliciting strong immune responses in both herpes simplex virus vaccine and pseudorabies virus vaccine. Some glycolipids have been synthesized from long chain-alkylamines and fatty acids that are linked directly with the sugars through the anomeric carbon atom, to mimic the functions of the naturally occurring lipid residues.

U.S. Patent No. 4,258,029 granted to Moloney, assigned to the assignee hereof and incorporated herein by reference thereto, teaches that octadecyl tyrosine hydrochloride (OTH) functions as an adjuvant when complexed with tetanus toxoid and formalin inactivated type I, II and III poliomyelitis virus vaccine. Also,

Nixon-George et al. 1990, (ref. 25) reported that octadecyl esters of aromatic amino acids complexed with a recombinant hepatitis B surface antigen, enhanced the host immune responses against hepatitis B virus.

5        **2. Immunoassays**

10        The transferrin receptor proteins, analogs and/or fragments thereof of the present invention are useful as immunogens, as antigens in immunoassays including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known in the art for the detection of anti-*Moraxella*, transferrin receptor protein antibodies. In ELISA assays, the transferrin receptor protein, analogs and/or fragments corresponding to portions of Tfr

15        protein, are immobilized onto a selected surface, for example, a surface capable of binding proteins or peptides such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed transferrin receptor, analogs and/or fragments, a non-specific protein such as a solution of bovine serum

20        albumin (BSA) or casein that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by non-specific bindings of antisera onto the

25        surface.

30        The immobilizing surface is then contacted with a sample, such as clinical or biological materials, to be tested in a manner conducive to immune complex (antigen/antibody) formation. This procedure may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to

35        incubate for from about 2 to 4 hours, at temperatures such as of the order of about 25° to 37°C. Following

incubation, the sample-contacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween or a borate buffer.

5        Following formation of specific immunocomplexes between the test sample and the bound transferrin receptor protein, analogs and/or fragments and subsequent washing, the occurrence, and even amount, of immunocomplex formation may be determined by subjecting  
10       the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and in general IgG. To provide detecting means, the second  
15       antibody may have an associated activity such as an enzymatic activity that will generate, for example, a color development upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for  
20       example, a spectrophotometer.

### 3. Use of Sequences as Hybridization Probes

      The nucleotide sequences of the present invention, comprising the sequence of the transferrin receptor gene, now allow for the identification and cloning of  
25       the transferrin receptor genes from any species of *Moraxella*.

      The nucleotide sequences comprising the sequence of the transferrin receptor genes of the present invention are useful for their ability to selectively form duplex  
30       molecules with complementary stretches of other TfR genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other TfR genes. For a high degree of selectivity,  
35       relatively stringent conditions are used to form the duplexes, such as low salt and/or high temperature

conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results. In general, convenient hybridization temperatures in the presence of 50% formamide are: 42°C for a probe which is 95 to 100% homologous to the target fragment, 37°C for 90 to 95% homology and 32°C for 85 to 90% homology.

In a clinical diagnostic embodiment, the nucleic acid sequences of the TfR genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin and digoxigenin-labelling, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with samples containing TfR gene sequences.

The nucleic acid sequences of TfR genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid-

phase procedures, the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e. g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the TfR genes or fragments thereof of the present invention under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required depending on, for example, the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label. It is preferred to select nucleic acid sequence portions which are conserved among species of *Moraxella*. The selected probe may be at least 18bp and may be in the range of about 30 to 90 bp.

#### 4. Expression of the Transferrin Receptor Genes

Plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell may be used for the expression of the transferrin receptor genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the host cell for

expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM<sup>TM</sup>-11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as *E. coli* LE392.

Promoters commonly used in recombinant DNA construction include the  $\beta$ -lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system as described in U.S. Patent No. 4,952,496. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with genes. The particular promoter used will generally be a matter of choice depending upon the desired results. Hosts that are appropriate for expression of the transferrin receptor genes, fragments, analogs or variants thereof, may include *E. coli*, *Bacillus* species, *Haemophilus*, fungi, yeast, *Moraxella*, *Bordetella*, or the baculovirus expression system may be used.

In accordance with this invention, it is preferred to make the transferrin receptor protein, fragment or analog thereof, by recombinant methods, particularly since the naturally occurring TfR protein as purified from a culture of a species of *Moraxella* may include trace amounts of toxic materials or other contaminants.

This problem can be avoided by using recombinantly produced TfR protein in heterologous systems which can be isolated from the host in a manner to minimize contaminants in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have LPS and are, therefore, endotoxin free. Such hosts include species of *Bacillus* and may be particularly useful for the

production of non-pyrogenic transferrin receptor, fragments or analogs thereof. Furthermore, recombinant methods of production permit the manufacture of Tbp1 or Tbp2 or respective analogs or fragments thereof, separate from one another which is distinct from the normal combined proteins present in *Moraxella*.

#### Biological Deposits

Certain vectors that contain at least a portion coding for a transferrin receptor protein from strains of *Moraxella catarrhalis* strain 4223 and Q8 and a strain of *M. catarrhalis* RH408 that are described and referred to herein have been deposited with the American Type Culture Collection (ATCC) located at 12301 Parklawn Drive, Rockville, Maryland, USA, pursuant to the Budapest Treaty and prior to the filing of this application. Samples of the deposited vectors and bacterial strain will become available to the public and the restrictions imposed on access to the deposits will be removed upon grant of a patent based upon this United States patent application. In addition, the deposit will be replaced if viable samples cannot be dispensed by the Depository. The invention described and claimed herein is not to be limited in scope by the biological materials deposited, since the deposited embodiment is intended only as an illustration of the invention. Any equivalent or similar vectors or strains that encode similar or equivalent antigens as described in this application are within the scope of the invention.



## Deposit Summary

DEPOSIT	ATCC DESIGNATION	DATE DEPOSITED
Phage LEM3-24	97,381	December 4, 1995
Phage SLRD-A	97,380	December 4, 1995
Plasmid pLEM29	97,461	March 8, 1996
Plasmid pSLRD35A	97,833	January 13, 1997
Plasmid pLEM37	97,834	January 13, 1997
Strain RH408	55,637	December 9, 1994

## EXAMPLES

5           The above disclosure generally describes the  
present invention. A more complete understanding can be  
obtained by reference to the following specific  
Examples. These Examples are described solely for  
10       purposes of illustration and are not intended to limit  
the scope of the invention. Changes in form and  
substitution of equivalents are contemplated as  
circumstances may suggest or render expedient. Although  
specific terms have been employed herein, such terms are  
15       intended in a descriptive sense and not for purposes of  
limitations.

      Methods of molecular genetics, protein biochemistry  
and immunology used but not explicitly described in this  
disclosure and these Examples are amply reported in the  
scientific literature and are well within the ability of  
20       those skilled in the art.

## Example 1

      This Example illustrates the preparation and  
immunization of guinea pigs with Tbp1 and Tbp2 proteins  
25       from *M. catarrhalis*.

Tbp1 and Tbp2 proteins were obtained as follows:

Iron-starved crude total membrane preparations were diluted to 4 mg protein/ml in 50 mM Tris.HCl-1M NaCl, pH 8, in a total volume of 384 ml. Membranes were solubilized by the addition of 8 ml each of 0.5M EDTA and 30% sarkosyl and samples were incubated for 2 hours at room temperature, with gentle agitation. Solubilized membranes were centrifuged at 10K rpm for 20 min. 15 ml of apo-hTf-Sepharose 4B were added to the supernatant, and incubated for 2 hours at room temperature, with gentle shaking. The mixture was added into a column. The column was washed with 50 ml of 50mM Tris.HCl-1 M NaCl-250mM guanidine hydrochloride, to remove contaminating proteins. Tbp2 was eluted from the column by the addition of 100 ml of 1.5M guanidine hydrochloride. Tbp1 was eluted by the addition of 100 ml of 3M guanidine hydrochloride. The first 20 ml fractions were dialyzed against 3 changes of 50 mM Tris.HCl, pH 8.0. Samples were stored at -20°C, or dialyzed against ammonium bicarbonate and lyophilized.

Guinea pigs (Charles River) were immunized intramuscularly on day +1 with a 10 µg dose of Tbp1 or Tbp2 emulsified in complete Freund's adjuvant. Animals were boosted on days +14 and +29 with the same dose of protein emulsified in incomplete Freund's adjuvant. Blood samples were taken on day +42, and sera were used for analysis of bactericidal antibody activity. In addition, all antisera were assessed by immunoblot analysis for reactivity with *M. catarrhalis* 4223 proteins.

The bactericidal antibody activity of guinea pig anti-*M. catarrhalis* 4223 Tbp1 or Tbp2 antisera was determined as follows. A non-clumping *M. catarrhalis* strain RH408, derived from isolate 4223, was inoculated into 20 ml of BHI broth, and grown for 18 hr at 37°C, shaking at 170 rpm. One ml of this culture was used to

inoculate 20 ml of BHI supplemented with 25 mM ethylenediamine-di-hydroxyphenylacetic acid (EDDA; Sigma). The culture was grown to an OD<sub>600</sub> of 0.5. The cells were diluted 1:200,000 in 140 mM NaCl, 93mM NaHCO<sub>3</sub>, 2mM Na barbiturate, 4mM barbituric acid, 0.5mM MgCl<sub>2</sub>.6H<sub>2</sub>O, 0.4mM CaCl<sub>2</sub>.2H<sub>2</sub>O, pH 7.6 (Veronal buffer), containing 0.1% bovine serum albumin (VBS) and placed on ice. Guinea pig anti-*M. catarrhalis* 4223 Tbp1 or Tbp2 antisera, along with prebleed control antisera, were heated to 56°C for 30 min. to inactivate endogenous complement. Serial twofold dilutions of each antisera in VBS were added to the wells of a 96-well Nunclon microtitre plate (Nunc, Roskilde, Denmark). Dilutions started at 1:8, and were prepared to a final volume of 25 µL in each well. 25 µL of diluted bacterial cells were added to each of the wells. A guinea pig complement (Biowhittaker, Walkersville, MD) was diluted 1:10 in VBS, and 25 µL portions were added to each well.

The plates were incubated at 37°C for 60 min, gently shaking at 70 rpm on a rotary platform. 50 µL of each reaction mixture were plated onto Mueller Hinton (Becton-Dickinson, Cockeysville, MD) agar plates. The plates were incubated at 37°C for 72 hr and the number of colonies per plate were counted. Bactericidal titres were assessed as the reciprocal of the highest dilution of antiserum capable of killing greater than 50% of bacteria compared with controls containing pre-immune sera. Results shown in Table 1 below illustrate the ability of the anti-Tbp1 and anti-Tbp2 guinea pig antisera to lyse *M. catarrhalis*.

#### Example 2

This Example illustrates the preparation of chromosomal DNA from *M. catarrhalis* strains 4223 and Q8.

*M. catarrhalis* isolate 4223 was inoculated into 100 ml of BHI broth, and incubated for 18 hr at 37°C with

shaking. The cells were harvested by centrifugation at 10,000 x g for 20 min. The pellet was used for extraction of *M. catarrhalis* 4223 chromosomal DNA.

The cell pellet was resuspended in 20 ml of 10 mM Tris-HCl (pH 7.5)-1.0 mM EDTA (TE). Pronase and SDS were added to final concentrations of 500 µg/ml and 1.0%, respectively, and the suspension was incubated at 37°C for 2 hr. After several sequential extractions with phenol, phenol:chloroform (1:1), and chloroform:isoamyl alcohol (24:1), the aqueous extract was dialysed, at 4°C, against 1.0 M NaCl for 4 hr, and against TE (pH 7.5) for a further 48 hr with three buffer changes. Two volumes of ethanol were added to the dialysate, and the DNA was spooled onto a glass rod. The DNA was allowed to air-dry, and was dissolved in 3.0 ml of water. Concentration was estimated, by UV spectrophotometry, to be about 290 µg/ml.

*M. catarrhalis* strain Q8 was grown in BHI broth as described in Example 1. Cells were pelleted from 50 ml of culture by centrifugation at 5000 rpm for 20 minutes, at 4°C. The cell pellet was resuspended in 10 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) and proteinase K and SDS were added to final concentrations of 500 µg/ml and 1%, respectively. The sample was incubated at 37°C for 4 hours until a clear lysate was obtained. The lysate was extracted twice with Tris-saturated phenol/chloroform (1:1), and twice with chloroform. The final aqueous phase was dialysed for 24 hours against 2 X 1000 ml of 1 M NaCl at 4°C, changing the buffer once, and for 24 hours against 2 x 1000 ml of TE at 4°C, changing the buffer once. The final dialysate was precipitated with two volume of 100% ethanol. The DNA was spooled, dried and resuspended in 5 to 10 ml of TE buffer.

### Example 3

This Example illustrates the construction of *M.*

*catarrhalis* chromosomal libraries in EMBL3.

5 A series of *Sau*3A restriction digests of chromosomal DNA, in final volumes of 10  $\mu$ L each, were carried out in order to optimize the conditions necessary to generate maximal amounts of restriction fragments within a 15 to 23 kb size range. Using the optimized digestion conditions, a large-scale digestion was set up in a 100  $\mu$ L volume, containing the following:

10 50  $\mu$ L of chromosomal DNA (290  $\mu$ g/ml), 33  $\mu$ L water, 10  $\mu$ L 10X *Sau*3A buffer (New England Biolabs), 1.0  $\mu$ L BSA (10 mg/ml, New England Biolabs), and 6.3  $\mu$ L *Sau*3A (0.04 U/ $\mu$ L). Following a 15 min. incubation at 37°C, the digestion was terminated by the addition of 10  $\mu$ L of 100 mM Tris-HCl (pH 8.0)-10 mM EDTA-0.1% bromophenol blue-50% glycerol (loading buffer). Digested DNA was electrophoresed through a 0.5% agarose gel in 40 mM Tris acetate-2 mM Na<sub>2</sub>EDTA.2H<sub>2</sub>O (pH 8.5) (TAE buffer) at 50 V for 6 hr. The region containing restriction fragments within a 15 to 23 kb molecular size range was excised from the gel, and placed into dialysis tubing containing 3.0 ml of TAE buffer. DNA was electroeluted from the gel fragment by applying a field strength of 1.0 V/cm for 18 hr. Electroeluted DNA was extracted once each with phenol and phenol:chloroform (1:1), and precipitated with ethanol. The dried DNA was dissolved in 5.0  $\mu$ L water.

20 Size-fractionated chromosomal DNA was ligated with *Bam*HI-digested EMBL3 arms (Promega), using T4 DNA ligase in a final volume of 9  $\mu$ L. The entire ligation mixture was packaged into lambda phage using a commercial packaging kit (Amersham), following manufacturer's instructions.

30 The packaged DNA library was amplified on solid media. 0.1 ml aliquots of *Escherichia coli* strain NM539 in 10 mM MgSO<sub>4</sub> (OD<sub>600</sub> = 0.5) were incubated at 37°C for 15

min. with 15 to 25  $\mu$ L of the packaged DNA library. Samples were mixed with 3 ml of 0.6% agarose containing 1.0% BBL trypticase peptone-0.5% NaCl (BBL top agarose), and mixtures were plated onto 1.5% agar plates containing 1.0% BBL trypticase peptone-0.5% NaCl, and incubated at 37°C for 18 hr. 3 ml quantities of 50 mM Tris-HCl (pH 7.5)-8 mM magnesium sulfate heptahydrate-100 mM NaCl-0.01% (w/v) gelatin (SM buffer) were added to each plate, and plates were left at 4°C for 7 hr. SM buffer containing phage was collected from the plates, pooled together, and stored in a screwcap tube at 4°C, with chloroform.

Chromosomal DNA from *M. catarrhalis* strain Q8 was digested with Sau3A I (0.1 unit/30  $\mu$ g DNA) at 37°C for 30 minutes and size-fractionated on a 0.6% low melting point agarose gel. DNA fragments of 15-23 kb were excised and the DNA was electroeluted for 25 minutes in dialysis tubing containing TAE (40 mM Tris acetate pH 8.5, 2 mM EDTA) at 150 V. The DNA was extracted once with phenol/chloroform (1:1), precipitated, and resuspended in water. The DNA was ligated overnight with EMBL3 BamH I arms (Promega) and the ligation mixture was packaged using the Lambda *in vitro* packaging kit (Stratagene) and plated onto *E. coli* LE392 cells. The library was titrated and stored at 4°C in the presence of 0.3% chloroform.

#### Example 4

This Example illustrates screening of the *M. catarrhalis* libraries.

Ten  $\mu$ L aliquots of phage stock from the EMBL3/4223 sample prepared in Example 3 above were combined each with 100  $\mu$ L of *E. coli* strain LE392 in 10 mM MgSO<sub>4</sub> (OD<sub>600</sub> = 0.5) (plating cells), and incubated at 37°C for 15 min.

The samples were mixed with 3 ml each of BBL top agarose, and the mixtures were poured onto 1.5% agarose

plates containing 1% bacto tryptone-0.5% bacto yeast extract-0.05% NaCl (LB agarose; Difco) and supplemented with 200  $\mu$ M EDDA. The plates were incubated at 37°C for 18 hr. Plaques were lifted onto nitrocellulose filters (Amersham Hybond-C Extra) using a standard protocol, and the filters were immersed into 5% bovine serum albumin (BSA; Boehringer) in 20 mM Tris-HCl (pH 7.5)-150 mM NaCl (TBS) for 30 min at room temperature, or 4°C overnight. Filters were incubated for at least 1 hr at room temperature, or 18 hr at 4°C, in TBS containing a 1/1000 dilution of guinea pig anti-*M. catarrhalis* 4223 Tbp1 antiserum. Following four sequential 10 min. washes in TBS with 0.05% Tween 20 (TBS-Tween), filters were incubated for 30 min. at room temperature in TBS-Tween containing a 1/4000 dilution of recombinant Protein G labelled with horseradish peroxidase (rProtein G-HRP; Zymed). Filters were washed as above, and submerged into CN/DAB substrate solution (Pierce). Color development was arrested by immersing the filters into water. Positive plaques were cored from the plates, and each placed into 0.5 ml of SM buffer containing a few drops of chloroform. The screening procedure was repeated two more times, until 100% of the lifted plaques were positive using the guinea pig anti-*M. catarrhalis* 4223 Tbp1 antiserum.

The EMBL3/Q8 library was plated onto LE392 cells on YT plates using 0.7% top agar in YT as overlay. Plaques were lifted onto nitrocellulose filters and the filters were probed with oligonucleotide probes labelled with  $^{32}$ P $\alpha$ -dCTP (Random Primed DNA labeling kit, Boehringer Mannheim). The pre-hybridization was performed in sodium chloride/sodium citrate (SSC) buffer (ref. 27) at 37°C for 1 hour and the hybridization was performed at 42°C overnight. The probes were based upon an internal sequence of 4223 *tbpA*:

I R D L T R Y D P G

(Seq ID No. 31)

4236-RD 5' ATTCGAGACTTAACACGCTATGACCCTGGC 3'

(Seq ID No 32)

5 4237-RD 5' ATTCGTGATTAACTCGCTATGACCCTGGT 3'

(Seq ID No 33).

Putative plaques were re-plated and submitted to second and third rounds of screening using the same procedures.

10 Phage clone SLRD-A was used to subclone the *tfr* genes for sequence analysis.

**Example 5**

15 This Example illustrates immunoblot analysis of the phage lysates using anti-*M. catarrhalis* 4223 Tbp1 and Tbp2 antisera.

Proteins expressed by the phage eluants selected in Example 4 above were precipitated as follows. 60 µL of each phage eluant were combined with 200 µL *E. coli* LE392 plating cells, and incubated at 37°C for 15 min.

20 The mixture was inoculated into 10 ml of 1.0% NZamine A-0.5% NaCl-0.1% casamino acids-0.5% yeast extract-0.2% magnesium sulfate heptahydrate (NZCYM broth), supplemented with 200 mM EDDA, and grown at 37°C for 18 hr, with shaking. DNase was added to 1.0 ml of the

25 culture, to a final concentration of 50 µg/ml, and the sample was incubated at 37°C for 30 min.

Trichloroacetic acid was added to a final concentration of 12.5%, and the mixture was left on ice for 15 min. Proteins were pelleted by centrifugation at 13,000 x g

30 for 10 min, and the pellet was washed with 1.0 ml of acetone. The pellet was air-dried and resuspended in 50 µL 4% SDS-20 mM Tris- HCl (pH 8.0)-0.2 mM EDTA (lysis buffer).

35 Following SDS-PAGE electrophoresis through an 11.5% gel, the proteins were transferred to Immobilon-P



filters (Millipore) at a constant voltage of 20 V for 18 hr, in 25mM Tris-HCl, 220mM glycine-20% methanol (transfer buffer). Membranes were blocked in 5% BSA in TBS for 30 min. at room temperature. Blots were exposed either to guinea pig anti-*M. catarrhalis* 4223 Tbp1, or to guinea pig anti-*M. catarrhalis* 4223 Tbp2 antiserum, diluted 1/500 in TBS-Tween, for 2 hr at room temperature. Following three sequential 10 min. washes in TBS-Tween, membranes were incubated in TBS-Tween containing a 1/4000 dilution of rProtein G-HRP for 30 min. at room temperature. Membranes were washed as described above, and immersed into CN/DAB substrate solution. Color development was arrested by immersing blots into water.

Three EMBL3 phage clones expressed both a 115 kDa protein which reacted with anti-Tbp1 antiserum, and an 80 kDa protein, which reacted with anti-Tbp2 antiserum on Western blots and were thus concluded to contain genes encoding the transferrin receptor proteins of *Moraxella catarrhalis*.

#### Example 6

This Example illustrates the subcloning of the *M. catarrhalis* 4223 Tbp1 protein gene, *tbpA*.

Plate lysate cultures of the recombinant phage described in Example 5 were prepared by combining phage eluant and *E. coli* LE392 plating cells, to produce confluent lysis on LB agar plates. Phage DNA was extracted from the plate lysates using a Wizard Lambda Preps DNA Purification System (Promega), according to manufacturer's instructions.

The EMBL3 clone LM3-24 was found to contain a 13.2 kb insert, flanked by two *SalI* sites. A probe to a *tbpA* gene was prepared and consisted of a 300 base pair amplified product generated by PCR using two degenerate oligonucleotide primers corresponding to an amino acid sequence of part of the Tbp1 protein (Figure 1). The

primer sequences were based upon the amino acid sequences NEVTGLG (SEQ ID No: 17) and GAINEIE (SEQ ID No: 18), which had been found to be conserved among the deduced amino acid sequences from several different *N. meningitidis* and *Haemophilus influenzae* *tbpA* genes. The amplified product was cloned into pCRII (Invitrogen, San Diego, CA) and sequenced. The deduced amino acid sequence shared homology with other putative amino acid sequences derived from *N. meningitidis* and *H. influenzae* *tbpA* genes (Figure 12). The subclone was linearized with *NotI* (New England Biolabs), and labelled using a digoxigenin random-labelling kit (Boehringer Mannheim), according to manufacturer's instructions. The concentration of the probe was estimated to be 2 ng/ $\mu$ L.

DNA from the phage clone was digested with *HindIII*, *AvrII*, *SalI/SphI*, or *SalI/AvrII*, and electrophoresed through a 0.8% agarose gel. DNA was transferred to a nylon membrane (Genescreen Plus, Dupont) using an LKB VacuGene XL vacuum transfer apparatus (Pharmacia). Following transfer, the blot was air-dried, and pre-hybridized in 5X SSC-0.1% N-lauroylsarcosine-0.02% sodium dodecyl sulfate-1.0% blocking reagent (Boehringer Mannheim) in 10 mM maleic acid-15 mM NaCl (pH 7.5) (pre-hybridization solution). Labelled probe was added to the pre-hybridization solution to a final concentration of 6 ng/ml, and the blot was incubated in the probe solution at 42°C for 18 hr. The blot was washed twice in 2X SSC-0.1% SDS, for 5 min. each at room temperature, then twice in 0.1X SSC-0.1% SDS for 15 min. each at 60°C. Following the washes, the membrane was equilibrated in 100mM maleic acid-150 mM NaCl (pH 7.5) (buffer 1) for 1 min, then left in 1.0% blocking reagent (Boehringer Mannheim) in buffer 1 (buffer 2) for 60 min, at room temperature. The blot was exposed to anti-DIG-alkaline phosphatase (Boehringer Mannheim) diluted 1/5000 in buffer 2, for 30 min. at room temperature.

Following two 15 min. washes in buffer 1, the blot was equilibrated in 100 mM Tris-HCl (pH 9.5), 100 mM NaCl, 50 mM MgCl<sub>2</sub> (buffer 3) for 2 min. The blot was wetted with Lumigen PPD substrate (Boehringer-Mannheim),  
5 diluted 1/100 in buffer 3, then wrapped in Saran wrap, and exposed to X-ray film for 30 min. The probe hybridized to a 3.8 kb *Hind*III-*Hind*III, a 2.0 kb *Avr*II-*Avr*II, and a 4.2 kb *Sal*I-*Sph*I fragment.

In order to subclone the 3.8 kb *Hind*III-*Hind*III  
10 fragment into pACYC177, phage DNA from the EMBL3 clone, and plasmid DNA from the vector pACYC177 (New England Biolabs), were digested with *Hind*III, and fractionated by electrophoresis on a 0.8% agarose gel. The 3.8 kb  
15 *Hind*III-*Hind*III phage DNA fragment, and the 3.9 kb *Hind*III-*Hind*III pACYC177 fragment, were excised from the gel and purified using a Geneclean kit (Bio 101 Inc., LaJolla, CA), according to manufacturer's directions. Purified insert and vector were ligated together using  
20 T4 DNA ligase (New England Biolabs), and transformed into *E. coli* HB101 (Gibco BRL). A Qiagen Plasmid Midi-Kit (Qiagen) was used to extract and purify sequencing-quality DNA from one of the ampicillin-resistant/kanamycin-sensitive transformants, which was  
25 found to carry a 3.8 kb *Hind*III-*Hind*III insert. The subclone was named pLEM3. As described in Example 7, below, subsequent sequencing revealed that pLEM3 contained the first about 2.0 kb of *tbpA* sequence (Figures 2 and 5).

In order to subclone the remaining 1 kb of the *tbpA*  
30 gene, a 1.6 kb *Hind*III-*Hind*III fragment was subcloned into pACYC177 as described above, and transformed by electroporation into *E. coli* HB101 (Gibco BRL). A Midi-Plasmid DNA kit (Qiagen) was used to extract plasmid DNA from a putative kanamycin-sensitive transformant  
35 carrying a plasmid with a 1.6 kb *Hind*III-*Hind*III insert. The subclone was termed pLEM25. As described in

Example 7 below, sequencing revealed that pLEM25 contained the remaining 1 kb of the *tbpA* gene (Figure 2 and 5).

#### Example 7

5 This Example illustrates the subcloning of the *M. catarrhalis* 4223 *tbpB* gene.

As described above, in all *Neisseriae* and *Haemophilus* species examined prior to the present invention, *tbpB* genes have been found immediately  
10 upstream of the *tbpA* genes which share homology with the *tbpA* gene of *M. catarrhalis* 4223. However, the sequence upstream of *M. catarrhalis* 4223 did not correspond with other sequences encoding *tbpB*.

In order to localize the *tbpB* gene within the EMBL3  
15 phage clone, a Southern blot was carried out using a degenerate probe from a highly conserved amino acid region within the Tbp2 protein. A degenerate oligonucleotide probe, was designed corresponding to the sequence encoding EGGFYGP (SEQ ID No: 30), which is  
20 conserved within the Tbp2 protein in a variety of *Neisseriae* and *Haemophilus* species. The probe was labelled with digoxigenin using an oligonucleotide tailing kit (Boehringer Mannheim), following the manufacturer's instructions. *Hind*III - digested EMBL3  
25 clone DNA was fractionated through a 0.8% agarose gel, and transferred to a Geneclean Plus nylon membrane as described in Example 6. Following hybridization as described above, the membrane was washed twice in 2X SSC-0.1% SDS, for 5 min. each at room temperature, then  
30 twice in 0.1X SSC-0.1% SDS for 15 min. each, at 50°C. Detection of the labelled probe was carried out as described above. The probe hybridized to a 5.5 kb *Nhe*I-*Sal*I fragment.

The 5.5 kb *Nhe*I-*Sal*I fragment was subcloned into  
35 pBR328 as follows. LEM3-24 DNA, and pBR328 DNA, were digested with *Nhe*I-*Sal*I, and electrophoresed through

0.8% agarose. The 5.5 kb *NheI*-*SalI* fragment, and the 4.9 kb pBR328 *NheI*-*SalI* fragments were excised from the gel, and purified using a Geneclean kit as described in Example 6. The fragments were ligated together using T4 DNA ligase, and transformed into *E. coli* DH5. A Midi-Plasmid DNA kit (Qiagen) was used to extract DNA from an ampicillin resistant / tetracycline sensitive clone containing a 5.5 kb *NheI*-*SalI* insert. This subclone was termed pLEM23. Sequencing revealed that pLEM23 contained 2 kb of the *tbpB* gene from *M. catarrhalis* 4223 (Figure 2).

**Example 8:**

This Example illustrates the subcloning of *M. catarrhalis* Q8 *tfr* genes.

The *M. catarrhalis* Q8 *tfr* genes were subcloned as follows. Phage DNA was prepared from plates. Briefly, the top agarose layer from three confluent plates was scraped into 9 ml of SM buffer (0.1 M NaCl, 0.2% MgSO<sub>4</sub>, 50 mM Tris-HCl, pH 7.6, 0.01% gelatin) and 100 µl of chloroform was added. The mixture was vortexed for 10 sec, then incubated at room temperature for 2h. The cell debris was removed by centrifugation at 8000 rpm for 15 min at 4°C in an SS34 rotor (Sorvall model RC5C). The phage was pelleted by centrifugation at 35,000 rpm in a 70.1 Ti rotor at 10°C for 2h (Beckman model L8-80) and was resuspended in 500 µl of SM buffer. The sample was incubated at 4°C overnight, then RNase and DNase were added to final concentrations of 40 µg/ml and 10 µg/ml, respectively and the mixture incubated at 37°C for 1h. To the mixture were added 10 µl of 0.5 M EDTA and 5 µl of 10% SDS and the sample was incubated at 6°C for 15 min. The mixture was extracted twice with phenol/chloroform (1:1) and twice with chloroform and the DNA was precipitated by the addition of 2.5 volumes of absolute ethanol.

A partial restriction map was generated and fragments were subcloned using the external Sal I sites from EMBL3 and internal AvrII or EcoR I sites as indicated in Figure 4. In order to facilitate the subcloning, plasmid pSKMA was constructed which introduces a novel multiple cloning site into pBluescript.SK (Stratagene). Oligonucleotides were used to introduce restriction sites for Mst II, Sfi I, and Avr II between the Sal I and Hind III sites of pBluescript.SK:

	Sal I	Cla I	Mst II	Sfi I	Avr II	HindIII
	↓	↓	↓	↓	↓	↓
15	4639-RD	5'	TCGACGGTAT	CGATGGCC	TTAG	GGGC CTAGGA 3'
	(SEQ ID No: 34)					
	4640-RD	3'	GCCATA	GCTACCGG	AATC	CCCG GATCCTTCGA
	(SEQ ID No: 35)					

20 Plasmid pSLRD1 contains a ~1.5 kb Sal I-Avr II fragment cloned into pSKMA; plasmids pSLRD2 and pSLRD4 contain ~2 kb and 4 kb AvrII-AvrII fragments cloned into pSKMA, respectively and contain the complete *tbpA* gene. Plasmid pSLRD3 contains a ~2.3 kb AvrII-EcoR I fragment cloned into pSKMA and plasmid SLRD5 is a 22.7 kb EcoRI - EcoRI fragment cloned into pSKMA. These two clones contain the complete *tbpB* gene (Figure 7).

#### Example 9

30 This Example illustrates sequencing of the *M. catarrhalis* *tbp* genes.

Both strands of the *tbp* genes subcloned according to Examples 6 to 8 were sequenced using an Applied Biosystems DNA sequencer. The sequences of the *M. catarrhalis* 4223 and Q8 *tbpA* genes are shown in Figures 5 and 10 respectively. A derived amino acid sequence was compared with other Tbp1 amino acid sequences, including

those of *Neisseriae meningitidis*, *Neisseriae gonorrhoeae*, and *Haemophilus influenzae* (Figure 12). The sequence of the *M. catarrhalis* 4223 and Q8 *tbpB* genes are shown in Figures 6 and 11 respectively. In order to obtain sequence from the putative beginning of the *tbpB* gene of *M. catarrhalis* 4223, sequence data were obtained directly from the clone LEM3-24 DNA. This sequence was verified by screening clone DS-1754-1. The sequence of the translated *tbpB* genes from *M. catarrhalis* 4223 and Q8 shared homology with deduced Tbp2 amino acid sequences of *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Haemophilus influenzae* (Figure 13).

#### Example 10

This Example illustrates the generation of an expression vector to produce recombinant Tbp1 protein. The construction scheme is shown in Figure 14.

Plasmid DNA from subclone pLEM3, prepared as described in Example 6, was digested with *Hind*III and *Bgl*I to generate a 1.84 kb *Bgl*I-*Hind*III fragment, containing approximately two-thirds of the *tbpA* gene. *Bam*HI was added to the digest to eliminate a comigrating 1.89kb *Bgl*I-*Hind*III vector fragment. In addition, plasmid DNA from the vector pT7-7 was digested with *Nde*I and *Hind*III. To create the beginning of the *tbpA* gene, an oligonucleotide was synthesized based upon the first 61 bases of the *tbpA* gene to the *Bgl*I site; an *Nde*I site was incorporated into the 5' end. Purified insert, vector and oligonucleotide were ligated together using T4 ligase (New England Biolabs), and transformed into *E. coli* DH5 $\alpha$ . DNA was purified from one of the 4.4 kb ampicillin-resistant transformants containing correct restriction sites (pLEM27).

Purified pLEM27 DNA was digested with *Hind*III, ligated to the 1.6 kb *Hind*III-*Hind*III insert fragment

of pLEM25 prepared as described in Example 6, and transformed into *E. coli* DH5 $\alpha$ . DNA was purified from an ampicillin-resistant transformant containing the correct restriction sites (pLEM29), and was transformed  
5 by electroporation into BL21 (DE3) (Novagen; Madison, WI) to produce *E. coli* pLEM29B-1.

A single isolated transformed colony was used to inoculate 100 ml of YT broth containing 100 $\mu$ g/ml ampicillin, and the culture was grown at 37°C overnight, shaking at 200 rpm. 200  $\mu$ l of the overnight  
10 culture were inoculated into 10 ml of YT broth containing 100 $\mu$ g/ml ampicillin, and the culture was grown at 37°C to an OD<sub>578</sub> of 0.35. The culture was induced by the addition of 30  $\mu$ l of 100 mM IPTG, and  
15 the culture was grown at 37°C for an additional 3 hours. One ml of culture was removed at the time of induction (t=0), and at t=1 hr and t=3 hrs. One ml samples were pelleted by centrifugation, and resuspended in 4%SDS-20 mM Tris.Cl, pH 8-200  $\mu$ M EDTA  
20 (lysis buffer). Samples were fractionated on an 11.5% SDS-PAGE gel, and transferred onto Immobilon filters (Amersham). Blots were developed using anti-Tbpl (*M. catarrhalis* 4223) antiserum, diluted 1:1000, as the primary antibody, and rproteinG conjugated with horseradish peroxidase (Zymed) as the secondary  
25 antibody. A chemiluminescent substrate (Lumiglo; Kirkegaard and Perry Laboratories, Gaithersburg, MD) was used for detection. Induced recombinant proteins were visible on the Coomassie-stained gels (Fig 15).  
30 The anti-Tbpl (4223) antiserum recognized the recombinant proteins on Western blots.

#### Example 11

This Example illustrates the extraction and purification of recombinant Tbpl of *M. catarrhalis* 4223.

35 Recombinant Tbpl protein, which is contained in inclusion bodies, was purified from *E. coli* cells.



expressing the *tbpA* gene (Example 10), by a procedure as shown in Figure 16. *E. coli* cells from a 500 ml culture, prepared as described in Example 10, were resuspended in 50 ml of 50 mM Tris-HCl, pH 8.0 containing 0.1 M NaCl and 5 mM AEBSF (protease inhibitor), and disrupted by sonication (3 x 10 min. 70% duty circle). The extract was centrifuged at 20,000 x g for 30 min. and the resultant supernatant which contained > 85% of the soluble proteins from *E. coli* was discarded.

The remaining pellet (Figure 16, PPT<sub>1</sub>) was further extracted in 50 ml of 50 mM Tris, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA. After centrifugation at 20,000 x g for 30 min., the supernatant containing residual soluble proteins and the majority of the membrane proteins was discarded.

The remaining pellet (Figure 16, PPT<sub>2</sub>) was further extracted in 50 ml of 50 mM Tris, pH 8.0 containing 2M urea and 5 mM dithiothreitol (DTT). After centrifugation at 20,000 x g for 30 min., the resultant pellet (Figure 16, PPT<sub>3</sub>) obtained after the above extraction contained the purified inclusion bodies.

The Tbp1 protein was solubilized from PPT<sub>3</sub> in 50 mM Tris, pH 8.0, containing 6 M guanidine hydrochloride and 5 mM DTT. After centrifugation, the resultant supernatant was further purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris, pH 8.0, containing 2M guanidine hydrochloride and 5 mM DTT. The fractions were analyzed by SDS-PAGE and those containing purified Tbp1 were pooled. Triton X-100 was added to the pooled Tbp1 fraction to a final concentration of 0.1%. The fraction was then dialyzed overnight at 4°C against 50 mM Tris, pH 8.0 and then centrifuged at 20,000 x g for 30 min. The protein remained soluble under these conditions and the purified Tbp1 was stored at -20° C. The purification procedure shown in Figure

16 produced Tbp1 protein that was at least 70% pure as determined by SDS-PAGE analysis (Figure 17).

#### Example 12

5 This Example illustrates the construction of an expression plasmid for rTbp2 of *M. catarrhalis* 4223 without a leader sequence.

The construction scheme for the plasmid expressing rTbp2 is shown in Figure 18. Oligonucleotides were used to construct the first approximately 58 bp of the  
10 *M. catarrhalis* 4223 *tbpB* gene encoding the mature protein. An *NdeI* site was incorporated into the 5' end of the oligonucleotides:

5' TATGTGTGGTGGCAGTGGTGGTTCAAATCCACCTGCTCCTACGCCCATT  
15 CCAAATG (SEQ ID NO: 36) 3'

3' ACACACCACCGTCACCACCAAGTTTAGGTGGACGAGGATGCGGGTAAGG  
TTTACGATC (SEQ ID NO: 37) 5'

An *NheI*-*ClaI* fragment, containing approximately 1kb of  
20 the *tbpB* gene from pLEM23, prepared as described in Example 7, was ligated to the above oligonucleotides and inserted into pT7-7 cut with *NdeI*-*ClaI*, generating pLEM31, which thus contains the 5'-half of *tbpB*. Oligonucleotides also were used to construct the last  
25 approximately 104 bp of the *tbpB* gene, from the *AvaII* site to the end of the gene. A *BamHI* site was incorporated into the 3' end of the oligonucleotides:

5' GTCCAAATGCAAACGAGATGGGCGGGTCATTTACACACAACGCCGATG  
30 ACAGCAAAGCCTCTGTGGTCTTTGGCACAAAAGACAACAAGAAGTTAAGTAGTA  
G (SEQ ID NO: 38) 3'

3' GTTTACGTTTGCTCTACCCGCCCAGTAAATGTGTGTTGCGGCTACTGTC  
GTTTCGGAGACACCAGAAACCGTGTCTTTCTGTTGTTCTTCAATTCATCATCCTAG  
35 (SEQ ID NO: 39) 5'

A *Cla*I-*Ava*II fragment from pLEM23, containing approximately 0.9 kb of the 3'-end of the *tbpB* gene, was ligated to the *Ava*II-*Bam*HI oligonucleotides, and inserted into pT7-7 cut with *Cla*I-*Bam*HI, generating pLEM32. The 1.0 kb *Nde*I-*Cla*I insert from pLEM31 and the 1.0 kb *Cla*I-*Bam*HI insert from pLEM32 were then inserted into pT7-7 cut with *Nde*I-*Bam*HI, generating pLEM33 which has a full-length *tbpB* gene under the direction of the T7 promoter.

DNA was purified from pLEM33 and transformed by electroporation into electrocompetent BL21(DE3) cells (Novagen; Madison, WI), to generate strain pLEM33B-1. Strain pLEM33B-1 was grown, and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and transferred to membranes suitable for immunoblotting. Blots were developed using anti-4223 Tbp2 antiserum, diluted 1:4000, as the primary antibody, and rprotein G conjugated with horseradish peroxidase (Zymed) as the secondary antibody. A chemiluminescent substrate (Lumiglo; Kirkegaard and Perry Laboratories, Gaithersburg, MD) was used for detection. Induced recombinant proteins were visible on the Coomassie blue-stained gels (Fig. 19). The anti-4223 Tbp2 antiserum recognized the recombinant proteins on Western blots.

#### Example 13

This Example illustrates the generation of an expression plasmid for rTbp2 of *M. catarrhalis* 4223 with a leader sequence.

The construction scheme is shown in Figure 18. Oligonucleotides containing the natural leader sequence of the *M. catarrhalis* 4223 *tbpB* gene were used to construct the first approximately 115 bp of the *tbpB* gene to the *Nhe*I site. An *Nde*I site was incorporated into the 5' end of the oligonucleotides:

5' TATGAAACACATTCCTTTAACCACACTGTGTGTGGCAATCTCTGCCGTC  
TTATTAACCGCTTGTGGTGGCAGTGGTGGTTCAAATCCACCTGCTCCTACGCCCAT  
TCCAAATG (SEQ ID NO: 40) 3'

5 3' ACTTTGTGTAAGGAAATTGGTGTGACACACACCGTTAGAGACGGCAGAA  
TAATTGGCGAACACCACCGTCACCACCAAGTTTAGGTGGACGAGGATGCGGGTAAG  
GTTTACGATC (SEQ ID NO: 41) 5'

10 The *NdeI*-*NheI* oligonucleotides were ligated to pLEM33  
cut with *NdeI*-*NheI*, generating pLEM37, which thus  
contains a full-length 4223 *tbpB* gene encoding the Tbp2  
protein with its leader sequence, driven by the T7  
promoter.

15 DNA from pLEM37 was purified and transformed by  
electroporation into electrocompetent BL21(DE3) cells  
(Novagen; Madison, WI), to generate strain pLEM37B-2.  
pLEM37B-2 was grown, and induced using IPTG, as  
described above in Example 10. Expressed proteins were  
20 resolved by SDS-PAGE and transferred to membranes  
suitable for immunoblotting. Blots were developed  
using anti-4223 Tbp2 antiserum, diluted 1:4000, as the  
primary antibody, and rprotein G conjugated with  
horseradish peroxidase (Zymed) as the secondary  
antibody. A chemiluminescent substrate (Lumiglo;  
25 Kirkegaard and Perry Laboratories, Gaithersburg, MD)  
was used for detection. Induced recombinant proteins  
were visible on Coomassie-blue stained gels (Fig. 21).  
The anti-4223 Tbp2 antiserum recognized the  
recombinant proteins on Western blots.

#### 30 **Example 14**

This Example illustrates the construction of an  
expression plasmid for rTbp2 of *M. catarrhalis* Q8  
without a leader sequence.

35 The construction scheme for rTbp2 is shown in  
Figure 20. The 5'-end of the *tbpB* gene of *M.*  
*catarrhalis* Q8 was PCR amplified from the Cys<sup>1</sup> codon of

the mature protein through the Bsm I restriction site.

An Nde I restriction site was introduced at the 5' end, for later cloning into pT7-7, and the final PCR fragment was 238 bp in length. The PCR primers are indicated below:

5' GAATTCCATATG TGT GGT GGG AGC TCT GGT GGT TTC AAT C  
3' 5247.RD (SEQ ID No: 42)

5' CCCATGGCAGGTTCTTGAATGCCTGAAACT 3' 5236.RD  
(SEQ ID No: 43).

The Q8 *tbpB* gene was subcloned in two fragments contained on plasmids SLRD3 and SLRD5, prepared as described in Example 8. Plasmid SLRD3-5 was constructed to contain the full-length *tbpB* gene by digesting SLRD5 with EcoR I and Dra I, which releases the 3'-end of *tbpB*, and inserting this ~ 619 bp fragment into SLRD3 which had been digested with EcoR I and Sma I. The 1.85 kb Bsm I-BamH I fragment from SLRD 3-5 was ligated with the 238 bp PCR fragment and inserted into pT7-7 that had been digested with Nde I and BamH I, generating plasmid SLRD35B. This plasmid thus contains the full-length *tbpB* gene without its leader sequence, under the direction of the T7 promoter. DNA from SLRD35B was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain SLRD35BD which was grown and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and the induced Tbp2 protein was clearly visible by Coomassie blue staining (Fig. 19).

#### Example 15

This Example illustrates the generation of an expression plasmid for rTbp2 of *M. catarrhalis* Q8 with

a leader sequence.

The construction scheme for the rTbp2 is shown in Figure 20. The 5'-end of the Q8 *tbpB* gene was PCR amplified from the ATG start codon to the Bsm I restriction site. An Nde I site was engineered at the 5'-end, to facilitate cloning into the pT7-7 expression vector, and the final PCR fragment was 295 bp. The PCR primers are indicated below:

10                   Nde I   K   H   I   P   L   T  
5' GAATTCCATATG AAA CAC ATT CCT TTA ACC 3'   5235.RD  
(SEQ ID No: 44)  
  
5' CCCATGGCAGGTTCTTGAATGCCTGAAACT           3'   5236.RD  
15 (SEQ ID No: 43).

SLRD3-5 (Example 14) was digested with Bsm I and BamH I, generating a 1.85 kb fragment, which was ligated with the 295bp PCR fragment and ligated into pT7-7 that had been digested with Nde I and BamH I. The resulting plasmid SLRD35A thus contains the full-length Q8 *tbpB* gene with its endogenous leader sequence under the control of the T7 promoter. DNA from SLRD35A was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain SLRD35AD which was grown and induced using IPTG, as described above in Example 10. Expressed proteins were resolved by SDS-PAGE and the induced Tbp2 protein was clearly visible by Coomassie blue staining (Fig. 19).

#### 30 **Example 16**

This Example illustrates the extraction and purification of rTbp2 of *M. catarrhalis* 4223 and Q8 from *E. coli*.

35 pLEM37B (4223) and SLRD35AD (Q8) transformants were grown to produce Tbp2 in inclusion bodies and then the Tbp2 was purified according to the scheme in Figure

22. *E. coli* cells from a 500 mL culture, were resuspended in 50 mL of 50 mM Tris-HCl, pH 8.0 containing 5 mM AEBSF (protease inhibitor), and disrupted by sonication (3 x 10 min, 70% duty circle).  
5 The extract was centrifuged at 20,000 x *g* for 30 min and the resultant supernatant which contained > 95% of the soluble proteins from *E. coli* was discarded.

The remaining pellet (PPT<sub>1</sub>) was further extracted in 50 mL of 50 mM Tris, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA. The mixture was stirred at 4°C  
10 for at least 2 hours and then centrifuged at 20,000 x *g* for 30 min and the supernatant containing residual soluble proteins and the majority of the membrane proteins was discarded.

The resultant pellet (PPT<sub>2</sub>) obtained after the above extraction contained the inclusion bodies. The Tbp2 protein was solubilized in 50 mM Tris, pH 8.0, containing 6 M guanidine and 5 mM DTT. After centrifugation, the resultant supernatant was further  
20 purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris, pH 8.0, containing 2 M guanidine and 5 mM DTT. The fractions were analyzed by SDS-PAGE and those containing purified Tbp2 were pooled. Triton X-100 was added to the pooled Tbp2  
25 fraction to a final concentration of 0.1%. The fraction was then dialyzed overnight at 4°C against PBS, and then centrifuged at 20,000 x *g* for 30 min. The protein remained soluble under these conditions and the purified Tbp2 was stored at -20°C. Figure 22 shows the  
30 SDS PAGE analysis of fractions of the purification process for rTbp2 from strain 4223 (Panel A) and strain Q8 (Panel B). The rTbp2 was at least 70% pure.

Groups of five BALB/c mice were injected three times subcutaneously (s.c.) on days 1, 29 and 43 with  
35 purified rTbp2 (0.3 mg to 10 mg) from *M. catarrhalis* strains 4223 and Q8 in the presence or absence of AlPO<sub>4</sub>.

(1.5 mg per dose). Blood samples were taken on days 14, 28, 42 and 56 for analysing the anti-rTbp2 antibody titers by EIAs.

Groups of two rabbits and two guinea pigs (Charles River, Quebec) were immunized intramuscularly (i.m.) on day 1 with a 5 mg dose of purified rTbp2 protein emulsified in complete Freund's adjuvant (CFA). Animals were boosted on days 14 and 29 with the same dose of protein emulsified in incomplete Freund's adjuvant (IFA). Blood samples were taken on day 42 for analysing anti-rTbp2 antibody titers and bactericidal activity. Table 2 below shows the bactericidal activity of antibodies raised to the recombinant transferrin binding proteins rTbp1 (4223), rTbp2 (4223) and rTbp2 (Q8), prepared as described in these Examples, against *M. catarrhalis* strains 4223 and Q8.

#### Example 17

This Example illustrates the binding of Tbp2 to human transferrin *in vitro*.

Transferrin-binding activity of Tbp2 was assessed according to the procedures of Schryvers and Lee (ref. 28) with modifications. Briefly, purified rTbp2 was subjected to discontinuous electrophoresis through 12.5% SDS-PAGE gels. The proteins were electrophoretically transferred to PVDF membrane and incubated with horseradish peroxidase-conjugated human transferrin (HRP-human transferrin, 1:50 dilution) (Jackson ImmunoResearch Labs Inc., Mississauga, Ontario) at 4°C for overnight. LumiGLO substrate (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) was used for chemiluminescent detection of HRP activity according to the manufacturer's instructions. Both 4223 rTbp2 and Q8 rTbp2 bind to human transferrin under these conditions, as shown in Figure 24.

#### Example 18

This Example illustrates antigenic conservation of



Tbp2 amongst *M. catarrhalis* strains.

Whole cell lysates of *M. catarrhalis* strains and *E. coli* strains expressing recombinant Tbp2 proteins were separated by SDS-PAGE and electrophoretically transferred to PVDF membrane. Guinea pig anti-4223 rTbp2 or anti-Q8 rTbp2 antisera were used as first antibody and alkaline phosphatase conjugated goat anti-guinea pig antibody was used as second antibody to detect Tbp2. *M. catarrhalis* strains 3, 56, 135, 585, 4223, 5191, 8185 and ATCC 25240 were tested and all showed specific reactivity with anti-4223 rTbp2 or anti-Q8 rTbp2 antibody (Figure 25).

Table 3 illustrates the ability of anti-rTbp2 antibodies from one *M. catarrhalis* strain to recognize native or recombinant protein from a homologous or heterologous *M. catarrhalis* strain.

#### Example 19

This Example illustrates PCR amplification of the *tbpB* gene from *M. catarrhalis* strain R1 and characterization of the amplified R1 *tbpB* gene.

Chromosomal DNA from *M. catarrhalis* strain R1 was prepared using standard techniques. The design of the oligonucleotide sense primer was based on a region approximately 274 bases upstream of the *M. catarrhalis* 4223 *tbpB* gene, and the antisense primer was based upon a region approximately 11 bases downstream of the end of 4223 *tbpB*. The following primers were used:

sense primer (4940): 5' GATATAAGCACGCCCTACTT 3'  
(SEQ ID No: 48)  
antisense primer (4967): 5' CCCATCAGCCAAACAAACATTGTGT 3'  
(SEQ ID No: 49).

Each reaction tube contained 10 mM Tris-HCl (pH 8.85), 25 mM KCl, 5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 800 mM dNTPs, 1.0 mg each of primers 4940 and 4967, 10 ng of R1 DNA, and 2.5 U Pwo DNA polymerase (Boehringer

Mannheim) in a total volume of 100  $\mu$ l. The thermocycler was programmed for 5 min at 95°C, followed by 25 cycles of 95°C for 30 sec, 50°C for 45 sec, and 72°C for 2 min, and a 10 min final elongation elongation at 72°C. The amplified product was purified using a Geneclean (BIO 101) according to the manufacturer's instructions, and sequenced.

A partial restriction map of *M. catarrhalis* strain R1 *tbpB* prepared as just described is shown in Figure 26. The nucleotide and deduced amino acid sequences of the PCR amplified R1 *tbpB* gene are shown in Figure 27. The R1 *tbpB* gene encodes a 714 amino acid protein of molecular weight 76.8 kDa. The leader sequence of the R1 Tbp2 protein is identical to that of the 4223 and Q8 Tbp2 proteins. When the deduced R1 Tbp2 sequence was aligned with the 4223 Tbp2 sequence, it was found to be 83% identical and 88% homologous (Fig. 28). The conserved LEGGFY (SEQ ID No: 50) epitope was present, as found in Tbp2 from other *M. catarrhalis* strains as well as the *H. influenzae* and *N. meningitidis* Tbp2 proteins.

#### SUMMARY OF THE DISCLOSURE

In summary of this disclosure, the present invention provides purified and isolated DNA molecules containing transferrin receptor genes of *Moraxella catarrhalis*, the sequences of these transferrin receptor genes, and the derived amino acid sequences thereof. The genes and DNA sequences are useful for diagnosis, immunization, and the generation of diagnostic and immunological reagents. Immunogenic compositions, including vaccines, based upon expressed recombinant Tbp1 and/or Tbp2, portions thereof, or analogs thereof, can be prepared for prevention of diseases caused by *Moraxella*. Modifications are possible within the scope of this invention.

**TABLE I****BACTERICIDAL ANTIBODY TITRES FOR  
M. CATARRHALIS ANTIGENS**

ANTIGEN <sup>1</sup>	SOURCE OF ANTISERA <sup>2</sup>	BACTERICIDAL TITRE <sup>3</sup> RH408 <sup>4</sup>		BACTERICIDAL TITRE Q8 <sup>5</sup>	
		Pre-Immune	Post-Immune	Pre-Immune	Post-Immune
TBP1	GP	< 3.0	4.2-6.9	< 3.0	4.4.-6.2
TBP2	GP	< 3.0	12.0-13.6	< 3.0	< 3.0-4.0

- 1 antigens isolated from *M. catarrhalis* 4223
- 2 GP = guinea pig
- 3 bactericidal titres: expressed in log<sub>2</sub> as the dilution of antiserum capable of killing 50% of cells
- 4 *M. catarrhalis* RH408 is a non-clumping derivative of *M. catarrhalis* 4223
- 5 *M. catarrhalis* Q8 is a clinical isolate which displays a non-clumping phenotype

RECTIFIED SHEET (RULE 91)  
ISA/EP

**TABLE 2**

Antigen	Bactericidal titre - RH408		Bactericidal titre - Q8	
	pre-immune	post-immune	pre-immune	post-immune
rTbp1 (4223)	< 3.0	< 3.0	< 3.0	< 3.0
rTbp2 (4223)	< 3.0	10-15	< 3.0	< 3.0
rTbp2 (Q8)	NT	NT	< 3.0	5.5-7.5

Antibody titres are expressed in  $\log_2$  as the dilution of antiserum capable of killing 50% of cells

NT = not tested

**TABLE 3**

ELISA titres for anti-rTbp2 antibodies recognizing native or rTbp2 from strain 4223 or rTbp2 from strain Q8

Coated antigen	Anti-rTbp2 (4223) Antibody Titres		Anti-rTbp2 (Q8) Antibody Titres	
	Rabbit antisera	Guinea pig antisera	Rabbit antisera	Guinea pig antisera
Native Tbp2 (4223)	409,600	1,638,400	25,600	51,200
	204,800	1,638,400	25,600	102,400
rTbp2 (4223)	409,600	1,638,400	102,400	204,800
	409,600	1,638,400	102,400	204,800
rTbp2 (Q8)	409,600	1,638,400	1,638,400	1,638,400
	102,400	1,638,400	409,600	1,638,400

## REFERENCES

1. Brorson, J-E., A. Axelsson, and S.E. Holm. 1976. Studies on *Branhamella catarrhalis* (*Neisseria catarrhalis*) with special reference to maxillary sinusitis. *Scan. J. Infect. Dis.* 8:151-155.
2. Catlin, B.W., 1990. *Branhamella catarrhalis*: an organism gaining respect as a pathogen. *Clin. Microbiol. Rev.* 3: 293-320.
3. Hager, H., A. Verghese, S. Alvarez, and S.L. Berk. 1987. *Branhamella catarrhalis* respiratory infections. *Rev. Infect. Dis.* 9:1140-1149.
4. McLeod, D.T., F. Ahmad, M.J. Croughan, and M.A. Calder. 1986. Bronchopulmonary infection due to *M. catarrhalis*. Clinical features and therapeutic response. *Drugs* 31(Suppl.3):109-112.
5. Nicotra, B., M. Rivera, J.I. Luman, and R.J. Wallace. 1986. *Branhamella catarrhalis* as a lower respiratory tract pathogen in patients with chronic lung disease. *Arch.Intern.Med.* 146:890-893.
6. Ninane, G., J. Joly, and M. Kraytman. 1978. Bronchopulmonary infection due to *Branhamella catarrhalis* 11 cases assessed by transtracheal puncture. *Br.Med.Jr.* 1:276-278.
7. Srinivasan, G., M.J. Raff, W.C. Templeton, S.J. Givens, R.C. Graves, and J.C. Mel. 1981. *Branhamella catarrhalis* pneumonia. Report of two cases and review of the literature. *Am.Rev. Respir. Dis.* 123:553-555.
8. West, M., S.L. Berk, and J.K. Smith. 1982. *Branhamella catarrhalis* pneumonia. *South.Med. J.* 75:1021-1023.
9. Christensen, J.J., and B. Bruun. 1985. Bacteremia caused by a beta-lactamase producing strain of *Branhamella catarrhalis*. *Acta.Pathol. Microbiol. Immunol. Scand. Sect.B* 93:273-275.
10. Craig, D.B., and P.A. Wehrle. 1983. *Branhamella catarrhalis* septic arthritis. *J. Rheumatol.* 10:985-986.

11. Guthrie, R., K. Bakenhaster, R. Nelson, and R. Woskobnick. 1988. *Branhamella catarrhalis* sepsis: a case report and review of the literature. *J. Infect. Dis.* 158:907-908.
12. Hiroshi, S., E.J. Anaissie, N. Khardori, and G.P. Bodey. 1988. *Branhamella catarrhalis* septicemia in patients with leukemia. *Cancer* 61:2315-2317.
13. O'Neill, J.H., and P.W. Mathieson. 1987. Meningitis due to *Branhamella catarrhalis*. *Aust. N.Z. J. Med.* 17:241-242.
14. Murphy, T.F. 1989. The surface of *Branhamella catarrhalis*: a systematic approach to the surface antigens of an emerging pathogen. *Pediatr. Infect. Dis. J.* 8:S75-S77.
15. Van Hare, G.F., P.A. Shurin, C.D. Marchant, N.A. Cartelli, C.E. Johnson, D. Fulton, S. Carlin, and C.H. Kim. Acute otitis media caused by *Branhamella catarrhalis*: biology and therapy. *Rev. Infect. Dis.* 9:16-27.
16. Jorgensen, J.H., Doern, G.V., Maher, L.A., Howell, A.W., and Redding, J.S., 1990 Antimicrobial resistance among respiratory isolates of *Haemophilus influenza*, *Moraxella catarrhalis*, and *Streptococcus pneumoniae* in the United States. *Antibicrob. Agents Chemother.* 34: 2075-2080.
17. Schryvers, A.B. and Morris, L.J. 1988 Identification and Characterization of the transferrin receptor from *Neisseria meningitidis*. *Mol. Microbiol.* 2:281-288.
18. Lee, B.C., Schryvers, A.B. Specificity of the lactoferrin and transferrin receptors in *Neisseria gonorrhoeae*. *Mol. Microbiol.* 1988; 2: 827-9.
19. Schryvers, A.B. Characterization of the human transferrin and lactoferrin receptors in *Haemophilus influenzae*. *Mol. Microbiol.* 1988; 2: 467-72.
20. Schryvers, A.B. and Lee, B.C. (1988) Comparative analysis of the transferrin and lactoferrin binding proteins in the family *Neisseriaceae*. *Can. J. Microbiol.* 35, 409-415.

21. Yu, R. and Schryvers, A.B., 1993. The interaction between human transferrin and transferrin binding protein 2 from *Moraxella* (*Branhamella*) *catarrhalis* differs from that of other human pathogens. *Microbiol. Pathogenesis*, 15:433-445.
22. O'Hagan, 1992. *Clin. Pharmacokinet.* 22:1
23. Ulmer et al., 1993. *Curr. Opin. Invest. Drugs* 2: 983-989.
24. Lockhoff, O., 1991. glycolipids as immunomodulators: Synthesis and properties. *Chem. Int. Ed. Engl.* 30: 1611-1620.
25. Nixon-George, 1990. *J. Immunol.* 14: 4798-4802.
26. Wallace, R.J. Jr., Nash, D.R., and Steingrube, V.A. 1990. Antibiotic susceptibilities and drug resistance in *Moraxella* (*Branhamella*) *catarrhalis*. *Am. J. Med.* 88 (5A): 465-50S.
27. F.M. Ausubel et al., Short protocols in Molecular Biology, Greene Publishing Associates and John Wiley and Sons.
28. Schryvers, A.B., Lee, B.C. 1989. Comparative analysis of the transferrin and lactoferrin binding proteins in the family *Neisseriaceae*. *Can. J. Microbiol.* 35: 409-415.



CLAIMS

What we claim is:

1. A purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein.
2. The nucleic acid molecule of claim 1 wherein the transferrin receptor protein is the transferrin receptor binding protein 1 (Tbp1) of the *Moraxella* strain.
3. The nucleic acid molecule of claim 2 wherein the transferrin receptor protein is the transferrin receptor binding protein 2 (Tbp2) of the *Moraxella* strain.
4. The nucleic acid molecule of claim 1 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.
5. The nucleic acid molecule of claim 4 wherein the strain of *Moraxella catarrhalis* is *Moraxella catarrhalis* 4223, Q8 or R1.
6. A purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:
  - (a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto;
  - (b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and
  - (c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b).
7. The nucleic acid molecule of claim 6, wherein the DNA sequence defined in (c) has at least about 90% sequence identity with any one of the DNA sequences

defined in (a) or (b).

8. The nucleic acid molecule of claim 6 wherein the DNA sequence defined in (c) is that encoding the equivalent transferrin receptor protein from another strain of *Moraxella*.

9. A vector adapted for transformation of a host comprising the nucleic acid molecule of claim 1 or 6.

10. The vector of claim 9 encoding a fragment of a transferrin receptor protein and having the characteristics of a plasmid selected from the group consisting of pLEM3, pLEM25, pLEM23, DS-1698-1-1, DS-1754-1, pSLRD2, pSLRD3, pSLRD4 and pSLRD5.

11. The vector of claim 9 further comprising expression means operatively coupled to the nucleic acid molecule for expression by the host of said transferrin receptor protein of a strain of *Moraxella* or the fragment or the analog of the transferrin receptor protein.

12. The vector of claim 11 having the characteristics of plasmid pLEM-29, pLEM-33, pLEM-37, SLRD35-A and SLRD35-B.

13. A transformed host containing an expression vector as claimed in claim 11.

14. A method of forming a substantially pure recombinant transferrin receptor protein of a strain of *Moraxella*, which comprises:

growing the transformed host of claim 13 to express a transferrin receptor protein as inclusion bodies,

purifying the inclusion bodies free from cellular material and soluble proteins,

solubilizing transferrin receptor protein from the purified inclusion bodies, and

purifying the transferrin receptor protein free

from other solubilized materials.

15. The method of claim 14 wherein said transferrin receptor protein comprises Tbp1 alone, Tbp2 alone or a mixture of Tbp1 and Tbp2.

16. The method of claim 15 wherein said transferrin receptor protein is at least about 70% pure.

17. The method of claim 16 wherein said transferrin receptor protein is at least about 90% pure.

18. A recombinant transferrin receptor protein or fragment or analog thereof producible by the transformed host of claim 12.

19. The protein of claim 18 which is transferrin receptor binding protein 1 (Tbp1) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

20. The protein of claim 18 which is transferrin receptor binding protein 2 (Tbp2) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

21. The protein of claim 18 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.

22. An immunogenic composition, comprising at least one active component selected from the group consisting of:

(A) a purified and isolated nucleic acid molecule encoding a transferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the transferrin receptor protein;

(B) a purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:

(a) a DNA sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 45 or 46) or the complementary DNA sequence thereto;

(b) a DNA sequence encoding an amino acid sequence as set out in Figure 5, 6, 10, 11 or 27 (SEQ ID Nos: 9, 10, 11, 12, 13, 14, 15, 16 or 47) or the complementary DNA sequence thereto; and

(c) a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b); or

(C) a recombinant transferrin receptor protein or fragment or analog thereof producible by a transformed host containing an expression vector comprising a nucleic acid molecule as defined in (A) or (B) and expression means operatively coupled to the nucleic acid molecule for expression by the host of the recombinant transferrin receptor protein or fragment or analog thereof;

and a pharmaceutically acceptable carrier therefor, said at least one active component producing an immune response when administered to a host.

23. A method for generating an immune response in a host, comprising administering to the host an immunoeffective amount of the immunogenic composition of claim 22.

24. A method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with the nucleic acid molecule of claim 1 or 6 to produce duplexes comprising the nucleic acid molecule and any said nucleic acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining production of the duplexes.

25. A diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

- (a) the nucleic acid molecule of claim 1 or 6;
- (b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any said nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and
- (c) means for determining production of the duplexes.

AMINO ACID SEQUENCES OF A CONSERVED PORTION OF  
Tbp1 PROTEIN FOR CONSTRUCTION OF DEGENERATE  
PRIMERS USED IN PCR AMPLIFICATION OF A PORTION  
OF THE *M. cattarhalis* 4223 *tbpA* GENE.

N E V T G L G

SEQ ID NO: 17

G A I N E I E

SEQ ID NO: 18

FIG.1

*M. catarrhalis* 4223 Transferrin Receptor Genes

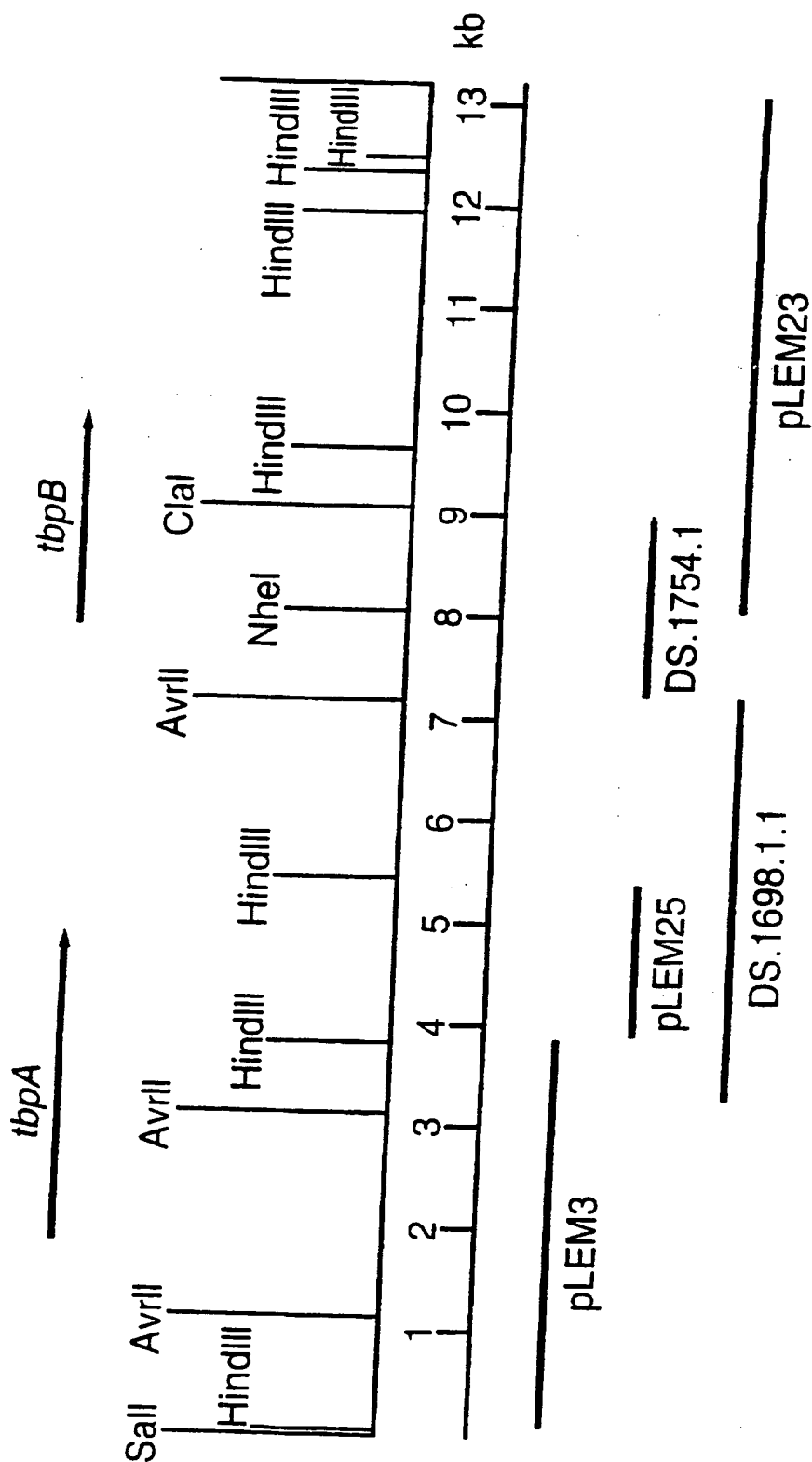


FIG.2

*M. catarrhalis* 4223 *tbpA* gene

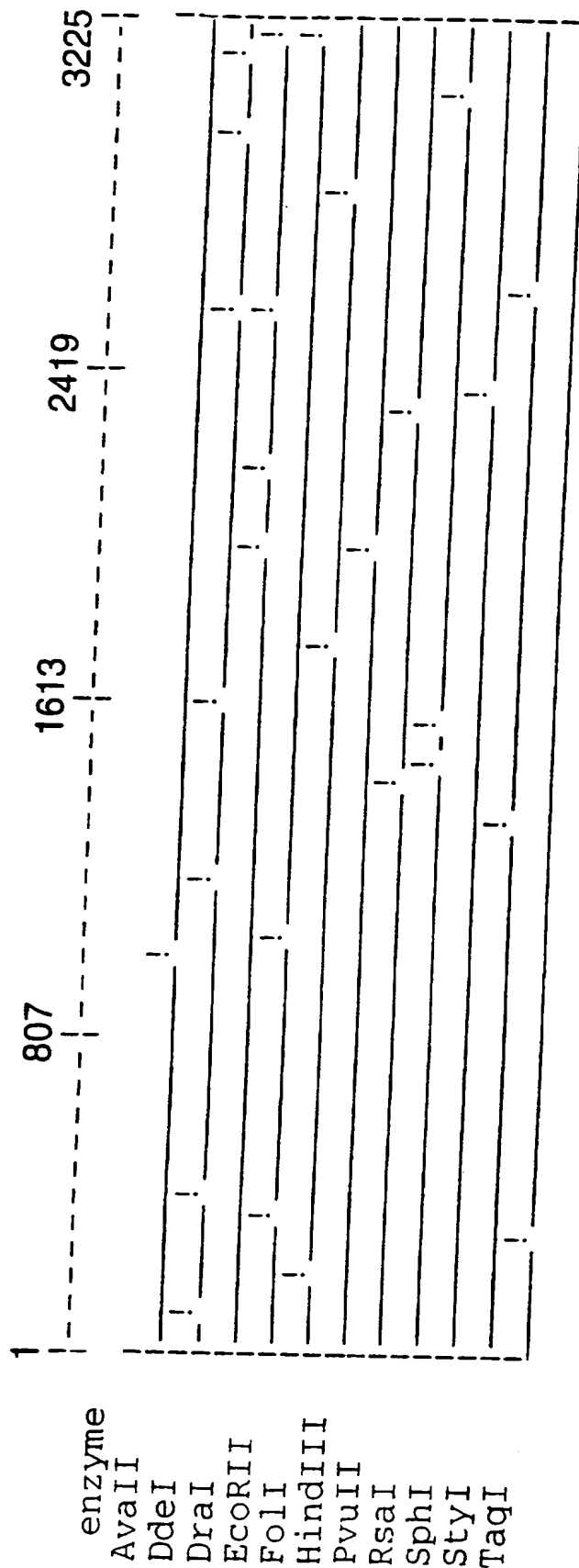


FIG.3



*M. catarrhalis* 4223 *tbpB* gene

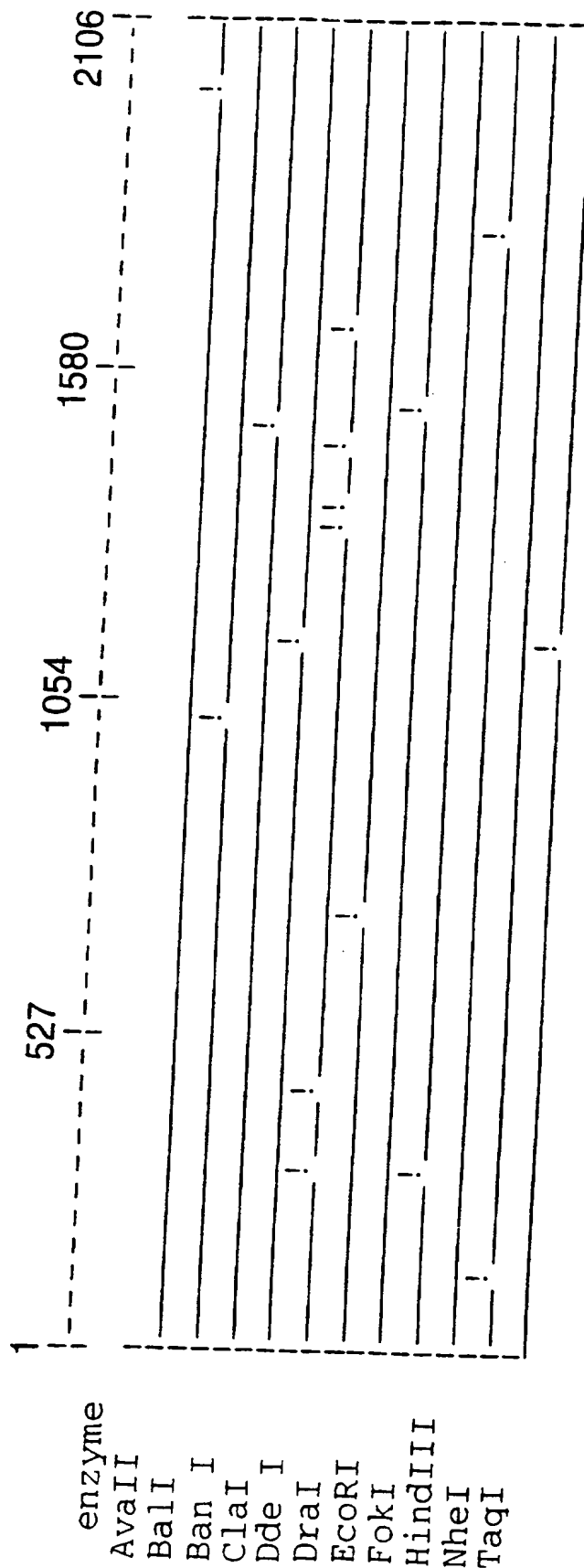


FIG.4

## FIG.5A

Sequence of *M. catarrhalis* 4223 *tbpA* gene

TATTTTGACAAGCTATACACTAAATCAAAATTAATCACTTTGGTTGGTGGTTTTAGCAAGCAAATGGT  
TATTTTGGTAAACAATTAAGTTCTTAAACGATACACGCTCATAAACAGATGGTTTGGCATCTGCAAT  
TTGATGCCCTGCCCTTGATGTTGGGTGTATCGGTGATCAAGTGCAAAAGCCACAGGTGGTCATTG

27  
ATG AAT CAA TCA AAA CAA AAC AAC AAA TCC AAA AAA TCC AAA CAA GTA TTA AAA  
MET Asn Gln Ser Lys Lys Gln Asn Asn Lys Ser Lys Lys Ser Lys Gln Val Leu Lys 54

81  
CTT AGT GCC TTG TCT TTG GGT CTG CTT AAC ATC ACG CAG GTG GCA CTG GCA AAC  
Leu Ser Ala Leu Ser Leu Gly Leu Leu Asn Ile Thr Gln Val Ala Leu Ala Asn 108

135  
ACA ACG GCC GAT AAG GCG GAG GCA ACA GAT AAG ACA AAC CTT GTT GTC TTTG  
Thr Thr Ala Asp Lys Lys Ala Glu Ala Thr Asp Lys Thr Asn Leu Val Val Leu 162

189  
GAT GAA ACT GTT GTA ACA GCG AAG AAA AAC GCC CGT AAA GCC AAC GAA GTT ACA  
Asp Glu Thr Val Val Thr Ala Lys Lys Asn Ala Arg Lys Ala Asn Glu Val Thr 216

FIG.5B

GGG CTT GGT AAG GTG GTC AAA ACT GCC	243	GAG ACC ATC AAT AAA GAA CAA GTG	270
Gly Leu Gly Lys Val Val Lys Thr Ala		Glu Thr Ile Asn Lys Glu Gln Val Leu	
AAC ATT CGA GAC TTA ACA CGC TAT	297	CCT GGC ATT GCT GTG GTT GAG CAA GGT	324
Asn Ile Arg Asp Leu Thr Arg Tyr Asp		Pro Gly Ile Ala Val Val Glu Gln Gly	
CGT GGG GCA AGC TCA GGC TAT TCT	351	CGT GGT ATG GAT AAA AAT CGT GTG GCG	378
Arg Gly Ala Ser Ser Gly Tyr Ser Ile		Arg Gly MET Asp Lys Asn Arg Val Ala	
GTA TTG GTT GAT GGC ATC AAT CAA GCC	405	CAG CAC TAT GCC CTA CAA GGC CCT GTG	432
Val Leu Val Val Asp Gly Ile Asn Gln Ala		Gln His Tyr Ala Leu Gln Gly Pro Val	
GCA GGC AAA AAT TAT GCC GCA GGT	459	GCA ATC AAC GAA ATA GAA TAC GAA AAT	486
Ala Gly Lys Lys Asn Tyr Ala Ala Gly		Ala Ile Asn Glu Ile Glu Tyr Glu Asn	
GTC CGC TCC GTT GAG ATT AGT AAA GGT	513	GCA AAT TCA AGT GAA TAC GGC TCT GGG	540
Val Arg Ser Val Glu Ile Ser Lys Gly		Ala Asn Ser Ser Glu Tyr Gly Ser Gly	

SUBSTITUTE SHEET (RULE 26)

## FIG.5C

GCA TTA TCT GGC TCT GTG GCA TTT GTT ACC AAA ACC GCC GAT GAC ATC ATC AAA	567	594
Ala Leu Ser Gly Ser Val Ala Phe Val Thr Lys Thr Ala Asp Asp Ile Ile Lys		
GAT GGT AAA GAT TGG GGC GTG CAG ACC AAA ACC GCC TAT GCC AGT AAA AAT AAC	621	648
Asp Gly Lys Asp Trp Gly Val Gln Thr Lys Thr Ala Tyr Ala Ser Lys Asn Asn		
GCA TGG GTT AAT TCT GTG GCA GCA GGC AAG GCA GGT TCT TTT AGC GGT CTT	675	702
Ala Trp Val Asn Ser Val Ala Ala Ala Gly Lys Ala Gly Ser Phe Ser Gly Leu		
ATC ATC TAC ACC GAC CGC CGT GGT CAA GAA TAC AAG GCA CAT GAT GAT GCC TAT	729	756
Ile Ile Tyr Thr Asp Arg Arg Gly Gln Glu Tyr Lys Ala His Asp Asp Ala Tyr		
CAG GGT AGC CAA AGT TTT GAT AGA GCG GTG GCA ACC ACT GAC CCA AAT AAC CGA	783	810
Gln Gly Ser Gln Ser Phe Asp Arg Ala Val Val Ala Thr Thr Asp Pro Asn Asn Arg		
ACA TTT TTA ATA GCA AAT GAA TGT GCC AAT GGT AAT TAT GAG GCG TGT GCT GCT	837	864
Thr Phe Leu Ile Ala Asn Glu Cys Ala Asn Gly Asn Tyr Glu Ala Cys Ala Ala		
GGC GGT CAA ACC AAA CTT CAA GCC AAG CCA ACC AAT GTG CGT GAT AAG GTC AAT	891	918
Gly Gly Gln Thr Lys Leu Gln Ala Lys Pro Thr Asn Val Arg Asp Lys Val Asn		

SUBSTITUTE SHEET (RULE 26)

## FIG.5D

GTC AAA GAT TAT ACA GGT CCT AAC CGC CTT ATC CCA AAC CCA CTC ACC CAA GAC	945	972
Val Lys Asp Tyr Thr Gly Pro Asn Arg Leu Ile Pro Asn Pro Leu Thr Gln Asp		
AGC AAA TCC TTA CTG CTT CGC CCA GGT TAT CAG CTA AAC GAT AAG CAC TAT GTC	999	1026
Ser Lys Ser Leu Leu Leu Arg Pro Gly Tyr Gln Leu Asn Asp Lys His Tyr Val		
GGT GGT GTG TAT GAA ATC ACC AAA CAA AAC TAC GCC ATG CAA GAT AAA ACC GTG	1053	1080
Gly Gly Val Tyr Glu Ile Thr Lys Gln Asn Tyr Ala MET Gln Asp Lys Thr Val		
CCT GCT TAT CTG ACG GTT CAT GAC ATT GAA AAA TCA AGG CTC AGC AAC CAT GCC	1107	1134
Pro Ala Tyr Leu Thr Val His Asp Ile Glu Lys Ser Arg Leu Ser Asn His Ala		
CAA GCC AAT GGC TAT TAT CAA GGC AAT AAT CTT GGT GAA CGC ATT CGT GAT ACC	1161	1188
Gln Ala Asn Gly Tyr Tyr Gln Gly Asn Asn Leu Gly Glu Arg Ile Arg Asp Thr		
ATT GGG CCA GAT TCA GGT TAT GGC ATC AAC TAT GCT CAT GGC GTA TTT TAT GAT	1215	1242
Ile Gly Pro Asp Ser Gly Tyr Gly Ile Asn Tyr Ala His Gly Val Phe Tyr Asp		

SUBSTITUTE SHEET (RULE 26)

## FIG.5E

GAA AAA CAC CAA AAA GAC CGC CTA GGG CTT GAA TAT GTT TAT GAC AGC AAA GGT	1269	1296
Glu Lys His Gln Lys Asp Arg Leu Gly Leu Glu Tyr Val Tyr Asp Ser Lys Gly		
GAA AAT AAA TGG TTT GAT GAT GTG CGT GTG TCT TAT GAT AAG CAA GAC ATT ACG	1323	1350
Glu Asn Lys Lys Trp Phe Asp Asp Val Arg Val Ser Tyr Asp Lys Gln Asp Ile Thr		
CTA CGC AGC CAG CTG ACC AAC ACG CAC TGT TCA ACC TAT CCG CAC ATT GAC AAA	1377	1404
Leu Arg Ser Gln Leu Thr Asn Thr His Cys Ser Thr Tyr Pro His Ile Asp Lys		
AAT TGT ACG CCT GAT GTC AAT AAA CCT TTT TCG GTA AAA GAG GTG GAT AAC AAT	1431	1458
Asn Cys Thr Pro Asp Val Asn Lys Pro Phe Ser Val Lys Glu Val Asp Asn Asn		
GCC TAC AAA GAA CAG CAC AAT TTA ATC AAA GCC GTC TTT AAC AAA AAA ATG GCG	1485	1512
Ala Tyr Lys Glu Gln His Asn Leu Ile Lys Ala Val Phe Asn Lys Lys MET Ala		
TTG GGC AGT ACG CAT CAT CAC ATC AAC CTG CAA GTT GGC TAT GAT AAA TTC AAT	1539	1566
Leu Gly Ser Thr His His His Ile Asn Leu Gln Val Gly Tyr Asp Lys Phe Asn		
TCA AGC CTG AGC CGT GAA GAT TAT CGT TTG GCA ACC CAT CAG TCT TAT CAA AAA	1593	1620
Ser Ser Leu Ser Arg Glu Asp Tyr Arg Leu Ala Thr His Gln Ser Tyr Gln Lys		

SUBSTITUTE SHEET (RULE 26)

## FIG.5F

CTT GAT TAC ACC CCA CCA AGT AAC CCT	TTG CCA GAT AAG TTT AAG CCC ATT TTA	1647	1674
Leu Asp Tyr Thr Pro Pro Ser Asn Pro	Leu Pro Asp Lys Phe Lys Pro Ile Leu		
GGT TCA AAC AAC AAA CCC ATT TGC CTT	GAT GCT TAT GGT CAT GAC CAT	1701	1728
Gly Ser Asn Asn Lys Pro Ile Cys Leu	Asp Ala Tyr Gly Tyr Gly His Asp His		
CCA CAG GCT TGT AAC GCC AAA AAC AGC	ACT TAT CAA AAT TTT GCC ATC AAA AAA	1755	1782
Pro Gln Ala Cys Asn Ala Lys Asn Ser	Thr Tyr Gln Asn Phe Ala Ile Lys Lys		
GGC ATA GAG CAA TAC AAC CAA AAA ACC	AAT ACC GAT AAG ATT GAT TAT CAA GCC	1809	1836
Gly Ile Glu Gln Tyr Asn Gln Lys Thr	Asn Thr Asp Lys Ile Asp Tyr Gln Ala		
ATC ATT GAC CAA TAT GAT AAA CAA AAC	CCC AAC AGC ACC CTA AAA CCC TTT GAG	1863	1890
Ile Ile Asp Gln Tyr Asp Lys Gln Asn	Pro Asn Ser Thr Leu Lys Pro Phe Glu		
AAA ATC AAA CAA AGT TTG GGG CAA GAA	AAA TAC AAC AAG ATA GAC GAA CTT GGC	1917	1944
Lys Ile Lys Gln Ser Leu Gly Gln Glu	Lys Tyr Asn Lys Ile Asp Glu Leu Gly		

SUBSTITUTE SHEET (RULE 26)

## FIG.5G

TTT AAA GCT TAT AAA GAT TTA CGC AAC	1971	GAA TGG GCG GGT TGG ACT AAT GAC AAC	1998
Phe Lys Ala Tyr Lys Asp Leu Arg Asn		Glu Trp Ala Gly Trp Thr Asn Asp Asn	
AGC CAA CAA AAT GCC AAT AAA GGC ACG	2025	GAT AAT ATC TAT CAG CCA AAT CAA GCA	2052
Ser Gln Gln Asn Ala Asn Lys Gly Thr		Asp Asn Ile Tyr Gln Pro Asn Gln Ala	
ACT GTG GTC AAA GAT GAC AAA TGT AAA	2079	TAT AGC GAG ACC AAC AGC TAT GCT GAT	2106
Thr Val Val Lys Asp Asp Lys Cys Lys		Tyr Ser Glu Thr Asn Ser Tyr Ala Asp	
TGC TCA ACC ACT CGC CAC ATC AGT GGT	2133	GAT AAT TAT TTC ATC GCT TTA AAA GAC	2160
Cys Ser Thr Thr Arg His Ile Ser Gly		Asp Asn Tyr Phe Ile Ala Leu Lys Asp	
AAC ATG ACC ATC AAT AAA TAT GTT GAT	2187	TTG GGG CTG GGT GCT CGC TAT GAC AGA	2214
Asn MET Thr Ile Asn Lys Tyr Val Asp		Leu Gly Leu Gly Ala Arg Tyr Asp Arg	
ATC AAA CAC AAA TCT GAT GTG CCT TTG	2241	GTA GAC AAC AGT GCC AGC AAC CAG CTG	2268
Ile Lys His Lys Ser Asp Val Pro Leu		Val Asp Asn Ser Ala Ser Asn Gln Leu	



## FIG.5H

TCT TGG AAT TTT GGC GTG GTC GTC AAG CCC ACC AAT TGG CTG GAC ATC GCT TAT	2295	2322
Ser Trp Asn Phe Gly Val Val Val Lys Pro Thr Asn Trp Leu Asp Ile Ala Tyr		
AGA AGC TCG CAA GGC TTT CGC ATG CCA AGT TTT TCT GAA ATG TAT GGC GAA CGC	2349	2376
Arg Ser Ser Gln Gly Phe Arg MET Pro Ser Phe Ser Glu MET Tyr Gly Glu Arg		
TTT GGC GTA ACC ATC GGT AAA GGC ACG CAA CAT GGC TGT AAG GGT CTT TAT TAC	2403	2430
Phe Gly Val Thr Ile Gly Lys Gly Thr Gln His Gly Cys Lys Gly Leu Tyr Tyr		
ATT TGT CAG CAG ACT GTC CAT CAA ACC AAG CTA AAA CCT GAA AAA TCC TTT AAC	2457	2484
Ile Cys Gln Gln Thr Val His Gln Thr Lys Leu Lys Pro Glu Lys Ser Phe Asn		
CAA GAA ATC GGA GCG ACT TTA CAT AAC CAC TTA GGC AGT CTT GAG GTT AGT TAT	2511	2538
Gln Glu Ile Gly Ala Thr Leu His Asn His Leu Gly Ser Leu Glu Val Ser Tyr		
TTT AAA AAT CGC TAT ACC GAT TTG ATT GTT GGT AAA AGT GAA GAG ATT AGA ACC	2565	2592
Phe Lys Asn Arg Tyr Thr Asp Leu Ile Val Gly Lys Ser Glu Glu Ile Arg Thr		
CTA ACC CAA GGT GAT AAT GCA GGC AAA CAG CGT GGT AAA GGT GAT TTG GGC TTT	2619	2646
Leu Thr Gln Gly Asp Asn Ala Gly Lys Gln Arg Gly Lys Gly Asp Leu Gly Phe		

SUBSTITUTE SHEET (RULE 26)

## FIG.5I

CAT AAT GGA CAA GAT GCT GAT TTG ACA	2673	GGC ATT AAC ATT CTT GGC AGA CTT GAC	2700
His Asn Gly Gln Asp Ala Asp Leu Thr		Gly Ile Asn Ile Leu Gly Arg Leu Asp	
CTA AAC GCT GTC AAT AGT CGC CTT CCC	2727	TAT GGA TTA TAC TCA ACA CTG GCT TAT	2754
Leu Asn Ala Val Asn Ser Arg Leu Pro		Tyr Gly Leu Tyr Ser Thr Leu Ala Tyr	
AAC AAA GTT GAT GTT AAA GGA AAA ACC	2781	TTA AAC CCA ACT TTG GCA GGA ACA AAC	2808
Asn Lys Val Asp Val Lys Gly Lys Thr		Leu Asn Pro Thr Leu Ala Gly Thr Asn	
ATA CTG TTT GAT GCC ATC CAG CCA TCT	2835	CGT TAT GTG GTG GGG CTT GGC TAT GAT	2862
Ile Leu Phe Asp Ala Ile Gln Pro Ser		Arg Tyr Val Val Gly Leu Gly Tyr Asp	
GCC CCA AGC CAA AAA TGG GGA GCA AAC	2889	GCC ATA TTT ACC CAT TCT GAT GCC AAA	2916
Ala Pro Ser Gln Lys Trp Gly Ala Asn		Ala Ile Phe Thr His Ser Asp Ala Lys	
AAT CCA AGC GAG CTT TTG GCA GAT AAG	2943	AAC TTA GGT AAT GGC AAC ATT CAA ACA	2970
Asn Pro Ser Glu Leu Leu Ala Asp Lys		Asn Leu Gly Asn Gly Asn Ile Gln Thr	

14/90

## FIG.5J

AAA CAA GCC ACC AAA GCA AAA TCC ACG	CCG TGG CAA ACA CTT GAT TTG TCA GGT	3024
Lys Gln Ala Thr Lys Ala Lys Ser Thr	Pro Trp Gln Thr Leu Asp Leu Ser Gly	
2997		
TAT GTA AAC ATA AAA GAT AAT TTT ACC	TTG CGT GCT GGC GTG TAC AAT GTA TTT	3078
Tyr Val Asn Ile Lys Asp Asn Phe Thr	Leu Arg Ala Gly Val Tyr Asn Val Phe	
3051		
AAT ACC TAT TAC ACC ACT TGG GAG GCT	TTA CGC CAA ACA GCA GAA GGG GCG GTC	3132
Asn Thr Tyr Tyr Thr Thr Trp Glu Ala	Leu Arg Gln Thr Ala Glu Gly Ala Val	
3105		
AAT CAG CAT ACA GGA CTG AGC CAA GAT	AAG CAT TAT GGT CGC TAT GCC GCT CCT	3186
Asn Gln His Thr Gly Leu Ser Gln Asp	Lys His Tyr Gly Arg Tyr Ala Ala Pro	
3159		
GGA CGC AAT TAC CAA TTG GCA CTT GAA	ATG AAG TTT TAA	
Gly Arg Asn Tyr Gln Leu Ala Leu Glu	MET Lys Phe	
3213		

## FIG.6A

Sequence of *M. catarrhalis* 4223 *tbpB* gene

GTAAATTGCCGTATTTTGTCTATCATAAATGCATTTATCAATGCTCAATAAATACGCCAAATGCACAT  
 TGTGAGCATGCCAAATAGGCATCAACAGACTTTTTTAGATAATACCATCAACCCATCAGAGGATTATTTT

27  
 ATG AAA CAC ATT CCT TTA ACC ACA CTG TGT GTG GCA ATC TCT GCC GTC TTA TTA 54  
 MET Lys His Ile Pro Leu Thr Thr Leu Cys Val Ala Ile Ser Ala Val Leu Leu

81  
 ACC GCT TGT GGT GGC AGT GGT GGT TCA AAT CCA CCT GCT CCT ACG CCC ATT CCA 108  
 Thr Ala Cys Gly Gly Ser Gly Gly Ser Asn Pro Pro Ala Pro Thr Pro Ile Pro

135  
 AAT GCT AGC GGT TCA GGT AAT ACT GGT AAC ACT GGT AAT GCT GGC GGT ACT GAT 162  
 Asn Ala Ser Gly Ser Gly Asn Thr Gly Asn Thr Gly Asn Ala Gly Gly Thr Asp

189  
 AAT ACA GCC AAT GCA GGT AAT ACA GGC GGT ACA AAC TCT GGT ACA GGC AGT GCC 216  
 Asn Thr Ala Asn Ala Gly Asn Thr Gly Gly Thr Asn Ser Gly Thr Gly Ser Ala

243  
 AAC ACA CCA GAG CCA AAA TAT CAA GAT GTA CCA ACT GAG AAA AAT GAA AAA GAT 270  
 Asn Thr Pro Glu Pro Lys Tyr Gln Asp Val Pro Thr Glu Lys Asn Glu Lys Asp

SUBSTITUTE SHEET (RULE 26)

FIG.6B

AAA GTT TCA TCC ATT CAA GAA CCT GCC	297	ATG GGT TAT GGC ATG GCT TTG AGT	324
Lys Val Ser Ser Ile Gln Glu Pro Ala		MET Gly Tyr Gly MET Ala Leu Ser	Lys
ATT AAT CTA CAC AAC CGA CAA GAC ACG	351	CCA TTA GAT GAA AAA AAT ATC ATT	378
Ile Asn Leu His Asn Arg Gln Asp Thr		Pro Leu Asp Glu Lys Asn Ile Ile Thr	
TTA GAC GGT AAA AAA CAA GTT GCA GAA	405	GGT AAA AAA TCG CCA TTG CCA TTT	432
Leu Asp Gly Lys Lys Gln Val Ala Glu		Gly Lys Ser Pro Leu Pro Phe Ser	
TTA GAT GTA GAA AAT AAA TTG CTT	459	GAT GGC TAT ATA GCA AAA ATG AAT	486
Leu Asp Val Glu Glu Lys Lys Leu Leu		Gly Tyr Ile Ala Lys MET Asn Val Ala	
GAT AAA AAT GCC ATT GGT GAC AGA ATT	513	AAG AAA GGT AAT AAA GAA ATC TCC	540
Asp Lys Asn Ala Ile Ile Gly Asp Arg Ile		Lys Lys Gly Asn Lys Glu Ile Ser Asp	
GAA GAA CTT GCC AAA CAA ATC AAA GAA	567	GCT GTG CGT AAA AGC CAT GAG TTT	594
Glu Glu Leu Ala Lys Lys Ile Lys Glu		Ala Val Arg Lys Ser His Glu Phe Gln	

SUBSTITUTE SHEET (RULE 26)

## FIG.6C

CAA GTA TTA TCA CTG GAA AAC AAA	ATT TTT CAT TCA AAT GAC GGA ACA ACC	621	648
Gln Val Leu Ser Ser Leu Glu Asn Lys	Ile Phe His Ser Asn Asp Gly Thr Thr		
AAA GCA ACC ACA CGA GAT TTA AAA TAT	GTT GAT TAT GGT TAC TAC TTG GCG AAT	675	702
Lys Ala Thr Thr Arg Asp Leu Lys Tyr	Val Asp Tyr Gly Tyr Tyr Leu Ala Asn		
GAT GGC AAT TAT CTA ACC GTC AAA ACA	GAC AAA CTT TGG AAT TTA GGC CCT GTG	729	756
Asp Gly Asn Tyr Tyr Leu Thr Val Lys Thr	Asp Lys Leu Trp Asn Leu Gly Pro Val		
GGT GGT GTG TTT TAT AAT GGC ACA ACG	ACC GCC AAA GAG TTG CCC ACA CAA GAT	783	810
Gly Gly Val Phe Tyr Asn Gly Thr Thr	Thr Ala Lys Glu Leu Pro Thr Gln Asp		
GCG GTC AAA TAT AAA GGA CAT TGG GAC	TTT ATG ACC GAT GTT GCC AAC AGA AGA	837	864
Ala Val Lys Tyr Lys Gly His Trp Asp	Phe <u>MET Thr Asp Val Ala Asn Arg Arg</u>		
AAC CGA TTT AGC GAA GTG AAA GAA AAC	TCT CAA GCA GGC TGG TAT TAT GGA GCA	891	918
<u>Asn Arg Phe Ser Glu Val Lys</u>	Ser Gln Ala Gly Trp Tyr Tyr Gly Ala		

## FIG.6D

TCT TCA AAA GAT GAA TAC AAC CGC TTA TTA ACT AAA GAA GAC TCT GCC CCT GAT	945	972
Ser Ser Lys Asp Glu Tyr Asn Arg Leu Leu Thr Lys Glu Asp Ser Ala Pro Asp		
GGT CAT AGC GGT GAA TAT GGC CAT AGC AGT GAG TTT ACT GTT AAT TTT AAG GAA	999	1026
Gly His Ser Gly Glu Tyr Gly His Ser Ser Ser Glu Phe Thr Val Asn Phe Lys Glu		
AAA AAA TTA ACA GGT AAG CTG TTT AGT AAC CTA CAA GAC CGC CAT AAG GGC AAT	1053	1080
Lys Lys Leu Thr Gly Lys Leu Phe Ser Ser Asn Leu Gln Asp Arg His Lys Gly Asn		
GTT ACA AAA ACC GAA CGC TAT GAC ATC GAT GCC AAT ATC CAC GGC AAC CGC TTC	1107	1134
Val Thr Lys Thr Glu Arg Tyr Asp Ile Asp Ala Asn Ile His Gly Asn Arg Phe		
CGT GGC AGT GCC ACC GCA AGC AAT AAA AAT GAC ACA AGC AAA CAC CCC TTT ACC	1161	1188
Arg Gly Ser Ala Thr Ala Ser Asn Lys Asn Asp Thr Ser Ser Lys His Pro Phe Thr		
AGT GAT GCC AAC AAT AGG CTA GAA GGT GGT TTT TAT GGG CCA AAA GGC GAG GAG	1215	1242
Ser Asp Ala Asn Asn Arg Arg Leu Glu Gly Gly Gly Phe Tyr Gly Pro Lys Gly Glu Glu		





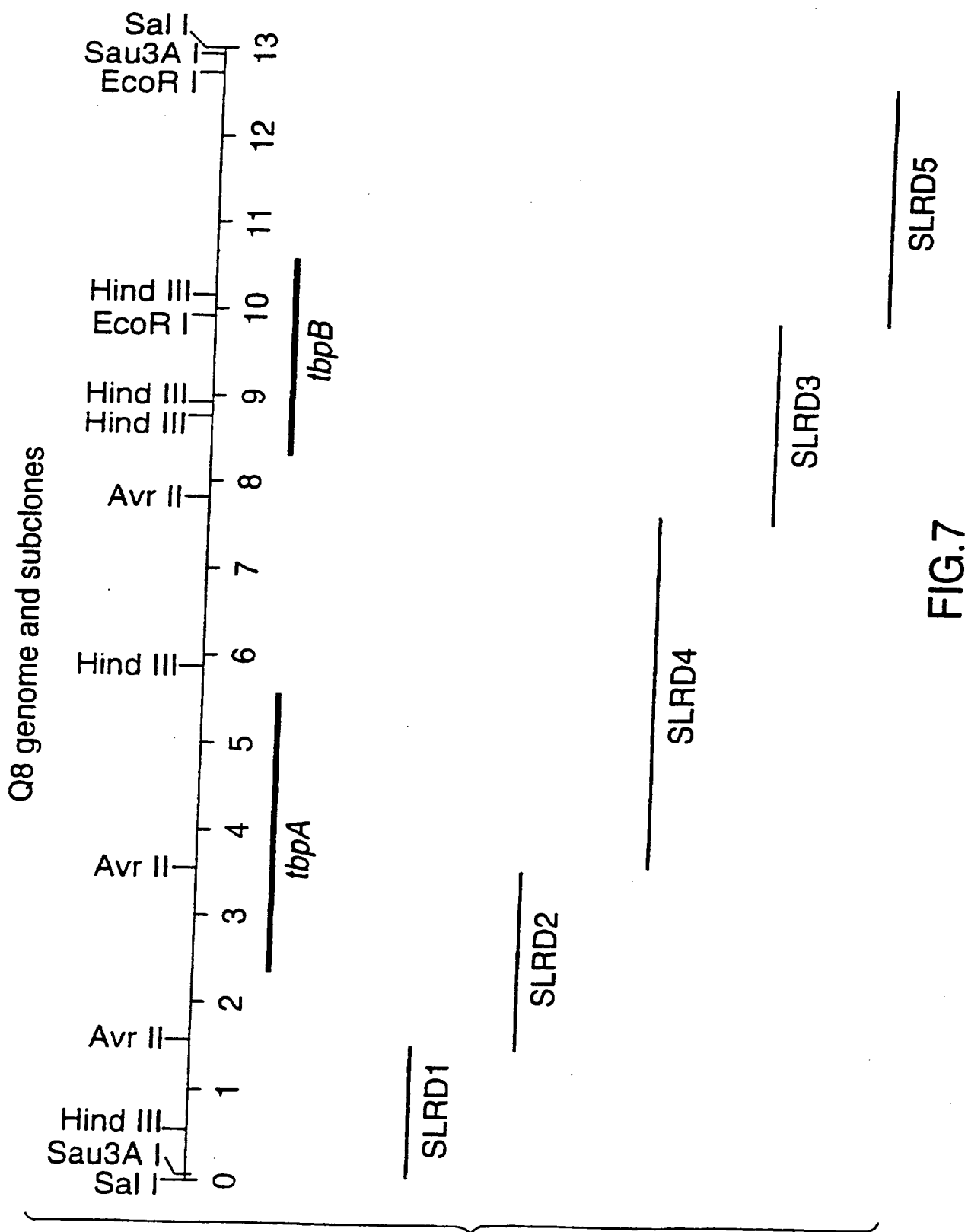
## FIG.6F

AAA ACC TAT GGC AAA AAC TTT GAA TAC CTA AAA TTT	1593	GGT GAG CTT AGT ATC GGT	1620
<u>Lys Thr Tyr Gly Lys Asn Phe Glu Tyr Leu Lys Phe</u>		<u>Gly Glu Leu Ser Ile Gly</u>	
GGT AGC CAT AGC GTC TTT TTA CAA GGC GAA CGC ACC GCT ACC ACA GGC GAG AAA	1647	Thr Gly Glu Lys	1674
Gly Ser His Ser Val Phe Leu Gln Gly Glu Arg Thr Ala Thr			
GCC GTA CCA ACC ACA GGC ACA GCC AAA TAT TTG GGG AAC TGG GTA GGA TAC ATC	1701	Tyr Leu Gly Asn Trp Val Gly Tyr Ile	1728
Ala Val Pro Thr Thr Gly Thr Ala Lys			
ACA GGA AAG GAC ACA GGA ACG GGC ACA GGA AAA AGC TTT ACC GAT GCC CAA GAT	1755	Thr Gly Lys Asp Thr Gly Thr Gly Lys Ser Phe Thr Asp Ala Gln Asp	1782
Val Gly Lys Asp Thr Gly Thr Gly Thr			
GTT GCT GAT TTT GAC ATT GAT TTT GGA AAT AAA TCA GTC AGC GGT AAA CTT ATC	1809	Lys Ser Val Ser Gly Lys Leu Ile	1836
Val Ala Asp Phe Asp Ile Asp Phe Gly Asn Lys Ser Val Ser			
ACC AAA GGC CGC CAA GAC CCT GTA TTT AGC ATC ACA GGT CAA ATC GCA GGC AAT	1863	Thr Lys Gly Arg Gln Asp Pro Val Phe Ser Ile Thr Gly Gln Ile Ala Gly Asn	1890

SUBSTITUTE SHEET (RULE 26)

## FIG.6G

GGC TGG ACA GGG ACA GCC AGC ACC ACC	AAA GCG GAC GCA GGA GGC TAC	1944
Gly Trp Thr Thr Thr Thr Thr Thr Thr	Lys Ala Asp Ala Gly Gly Tyr Lys	Ile
GAT TCT AGC AGT ACA GGC AAA TCC ATC	GCC ATC AAA GAT GCC AAT GTT ACA	1998
Asp Ser Ser Ser Thr Thr Gly Lys Ser Ile	Ala Ile Lys Asp Ala Asn Val Thr Gly	
GGC TTT TAT GGT CCA AAT GCA AAC GAG	ATG GGC GGC TCA TTT ACA CAC AAC	2052
Gly Phe Tyr Gly Gly Pro Asn Ala Asn Glu	MET Gly Gly Ser Phe Thr His Asn Ala	
GAT GAC AGC AAA GCC TCT GTG GTC TTT	GGC ACA AAA AGA CAA CAA GAA GTT	2106
Asp Asp Ser Lys Ala Ser Val Val Phe	Gly Thr Thr Lys Arg Gln Gln Val Lys	



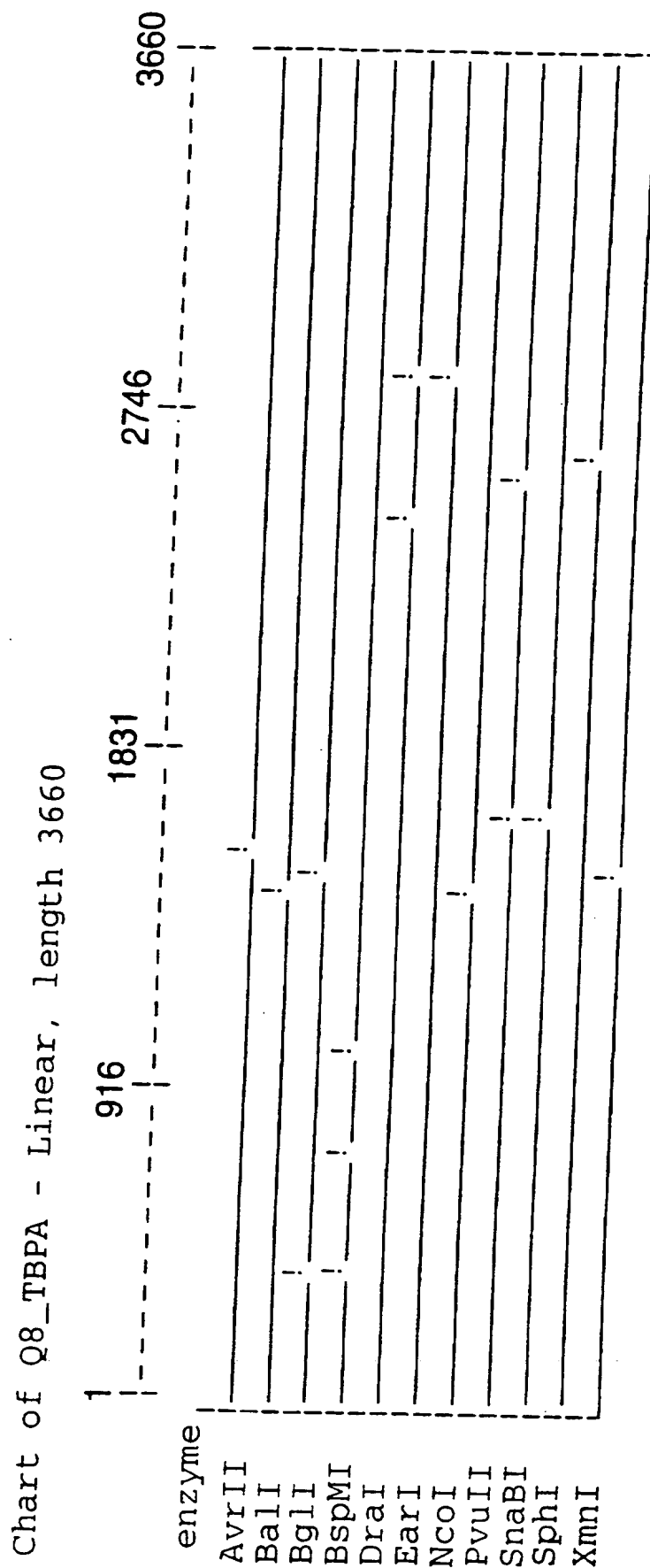


FIG.8

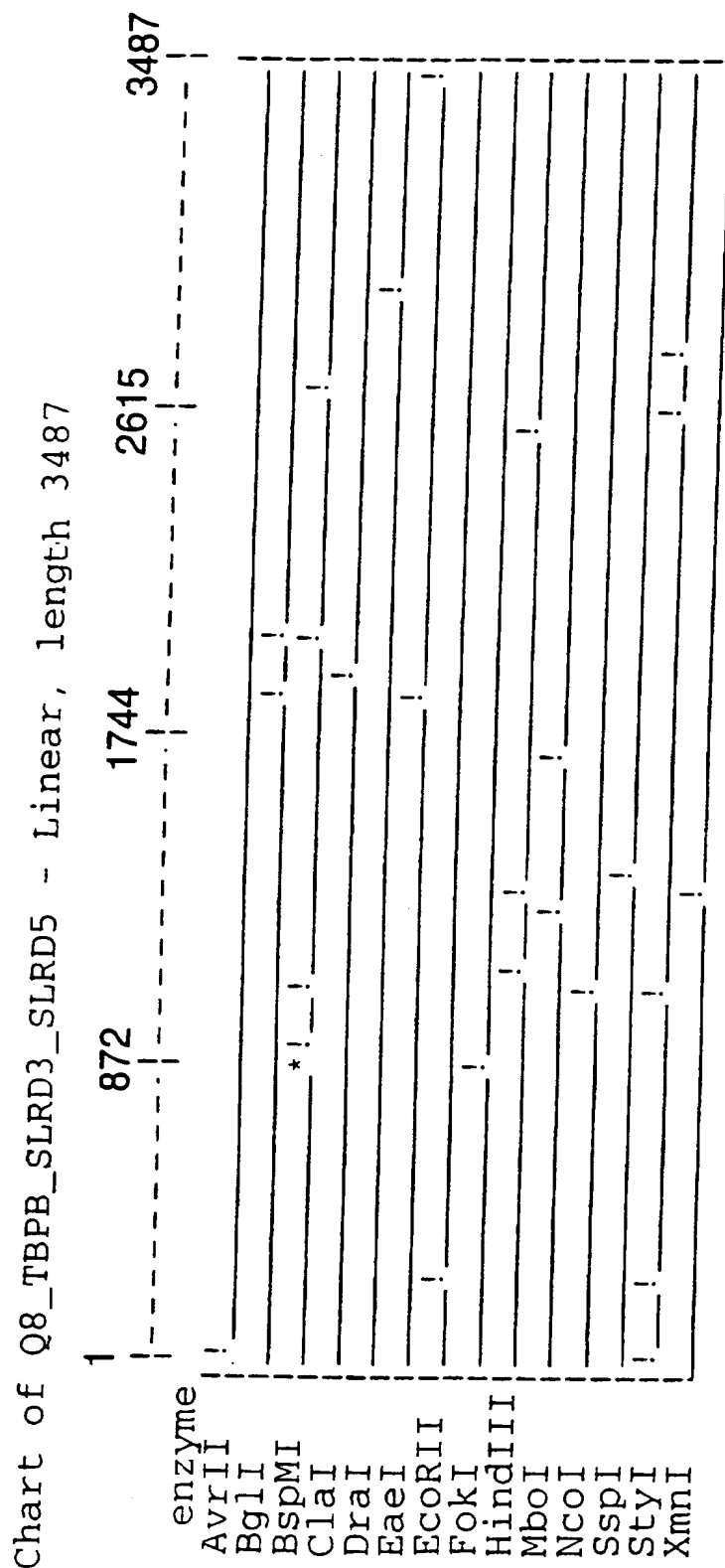


FIG.9

## FIG. 10A

Q8 tlpA gene sequence

```

A A T G A T A C A A A T G G T T T G T A T T A T C A C T      30
T G T A T T T G T A T T A T A A T T T T A C T T A T T T T      60
A C A A A C T A T A C A C T A A A A T C A A A A T T A A T      90
C A C T T T G G T T G G G T G G T T T T A G C A A G C A A A      120
T G G T T A T T T T G G T A A A C A A T T A A G T T C T T A      150
A A A A C G A T A C A C G C T C A T A A A C A G A T G G T T      180
T T T G G C A T C T T C A A T T T G A T G C C C T G C C T T G      210
T G A T T G G T T G G G G T G T A T T G A T G T A T C C A      240
A G T A C A A A G C C C A A C A G G T G G T C A T T G A T G      270
MET
```

SUBSTITUTE SHEET (RULE 26)

FIG.10B

ASN GLN SER LYS LYS SER LYS LYS SER LYS  
 AATCAATCCCAAATAATCCAAATAATCCAAA  
 280 290 300

GLN VAL LEU LYS LEU SER ALA LEU SER LEU  
 CAGTATTAAACTTAGTGCCCTTGCTTTG  
 310 320 330  
 GLY LEU LEU ASN ILE THR GLN VAL ALA LEU  
 GGTCTGCTTAACATCACGCAGGTGGCACTG  
 340 350 360

ALA ASN THR THR ALA ASP LYS ALA GLU ALA  
 GCAACACACGCGCGATTAAGCGGAGGCA  
 370 380 390  
 THR ASP LYS THR ASN LEU VAL VAL VAL LEU  
 ACAGATAAGACAAACCTTGTTGTTGTTGTTG  
 400 410 420

ASP GLU THR VAL VAL THR ALA LYS LYS ASN  
 GATGAACCTGTTGTACACGCGAAGAAAC  
 430 440 450  
 ALA ARG LYS ALA ASN GLU VAL THR GLY LEU  
 GCCCGTAAGCCCAACGAAGTTACAGGGCTT  
 460 470 480

FIG.10C

```

GLY  LYS  VAL  VAL  LYS  THR  ALA  GLU  THR  ILE
GGTAAGGTGGTCAAAACTGCCCAGACCATC
490
ASN  LYS  GLU  GLN  VAL  LEU  ASN  ILE  ARG  ASP
AATAAGAACAAAGTGCTAAACATTCCGAGAC
520
LEU  THR  ARG  TYR  ASP  PRO  GLY  ILE  ALA  VAL
TTACACGGCTATGACCCCTGGCATTTGCTGTG
550
VAL  GLU  GLN  GLY  ARG  GLY  ALA  SER  SER  GLY
GTTGAGCAAGGTCGTTGGGCAAGCTCAGGC
580
TYR  SER  ILE  ARG  GLY  MET  ASP  LYS  ASN  ARG
TATTCATTCTCGTGGTATGGATAAATACTCGT
610
VAL  ALA  VAL  LEU  VAL  ASP  GLY  ILE  ASN  GLN
GTGGCGGTATTGGTTGATGGCATCAATCAA
640
ALA  GLN  HIS  TYR  ALA  LEU  GLN  PRO  VAL
GCCAGCACTATGCCCTACAAAGGCCCTGTG
670

```

SUBSTITUTE SHEET (RULE 26)



## FIG.10D

```

ALA  GLY  LYS  ASN  TYR  ALA  ALA  GLY  GLY  ALA
G C A G G C A A A A T T A T G C C C G C A G G T G G G C C A
700                                710                                720

ILE  ASN  GLU  ILE  GLU  TYR  GLU  ASN  VAL  ARG
A T C A A C G A A A T A G A A T A C G A A A T G T C C G C
730                                740                                750

SER  VAL  GLU  ILE  SER  LYS  GLY  ALA  ASN  SER
T C C G T T G A G A T T A G T A A A G G T G C A A A T T C A
760                                770                                780

SER  GLU  TYR  GLY  SER  GLY  ALA  LEU  SER  GLY
A G T G A A T A C G G C T C T G G G G C A T T A T C T G G C
790                                800                                810

SER  VAL  ALA  PHE  VAL  THR  LYS  THR  ALA  ASP
T C T G T G G C A T T T G T T A C C A A A A C C G C C G A T
820                                830                                840

ASP  ILE  ILE  LYS  ASP  GLY  LYS  ASP  TRP  GLY
G A C A T C A A A G A T G G T A A A G A T T G G G C C
850                                860                                870

VAL  GLN  THR  LYS  THR  ALA  TYR  ALA  SER  LYS
G T G C A G A C C A A A C C G C C T A T G C C A G T A A A
880                                890                                900

```

SUBSTITUTE SHEET (RULE 26)

ASN ASN ALA TRP VAL ASN SER VAL ALA ALA  
AATAACGCATGGGTTAATTCTGTGGCAGCA  
910 920 930  
ALA GLY LYS ALA GLY SER PHE SER GLY LEU  
GCAGGCAAGGCAGGTCTTTTAGCGGTCCTT  
940 950 960  
ILE ILE TYR THR ASP ARG ARG GLY GLN GLU  
ATCATCTACACCGACCGCGTGGTCAAGAA  
970 980 990  
TYR LYS ALA HIS ASP ASP ATGATGATCATCAGGGT  
TACAAGGCACATGATGATGCCCTATCAGGGT  
1000 1010 1020  
SER GLN SER PHE ASP ARG ALA VAL ALA THR  
AGCCAAAGTTTGTATAGAGCGGTGGCAACC  
1030 1040 1050  
THR ASP PRO ASN ASN PRO LYS PHE LEU ILE  
ACTGACCCCAAATAACCCAAATAATTATAATA  
1060 1070 1080  
ALA ASN GLU CYS ALA ASN GLY ASN TYR GLU  
GCAATAATAATGTGCCAATGGTAATTATGAG  
1090 1100 1110

FIG. 10F

ALA CYS ALA ALA GLY GLN THR LYS LEU  
GCGTGCTGCTGCTGGCGGTCAACCAAACTC  
1120 1130 1140

GLN ALA LYS PRO THR ASN VAL ARG ASP LYS  
CAGCTAAGCCAAACCAATGTGCGTGATAAG  
1150 1160 1170

VAL ASN VAL LYS ASP TYR THR GLY PRO ASN  
GTCAATGTCAAAGATTATACAGGTCCCTAAC  
1180 1190 1200

ARG LEU ILE PRO ASN PRO LEU THR GLN ASP  
CGCCTTATCCCAACCCACTCAACCAGAC  
1210 1220 1230

SER LYS SER LEU LEU ARG PRO GLY TYR  
AGCAAATCCTTACTGCTTCGCCCAAGGTAT  
1240 1250 1260

GLN LEU ASN ASP LYS HIS TYR VAL GLY GLY  
CAGCTAACGATAAGCACTATGTCGGTGGT  
1270 1280 1290

VAL TYR GLU ILE THR LYS GLN ASN TYR ALA  
GTGTATGAATCACCACCAACAACTACGCC  
1300 1310 1320

## FIG.10G

```

MET  GLN  ASP  LYS  THR  VAL  PRO  ALA  TYR  LEU
A T G C A A G A T A A A C C G T G C C T T A T C T G
1330                                     1340
                                     THR  VAL  HIS  ASP  ILE  GLU  LYS  SER  ARG  LEU
                                     A C G G T T C A T G A C A T T G A A A A T C A A G G C T C
                                     1350                                     1360                                     1370                                     1380

SER  ASN  HIS  GLY  GLN  ALA  ASN  GLY  TYR  TYR
A G C A A C C A T G G C C A A G C C A A T G G C T A T T A T
1390                                     1400
                                     GLN  GLY  ASN  ASN  LEU  GLY  GLU  ARG  ILE  ARG
                                     C A A G G C A A T A A C C T T G G T G A A C G C A T T C G T
                                     1410                                     1420                                     1430                                     1440

ASP  ALA  ILE  GLY  ALA  ASN  SER  GLY  TYR  GLY
G A T G C C A T T G G G C A A A T T C A G G T T A T G G C
1450                                     1460                                     1470
                                     ILE  ASN  TYR  ALA  HIS  GLY  VAL  PHE  TYR  ASP
                                     A T C A A C T A T G C T C A T G G C G T A T T T A T G A C
                                     1480                                     1490                                     1500

GLU  LYS  HIS  GIN  LYS  ASP  ARG  LEU  GLY  LEU
G A A A A C A C C A A A A G A C C G C C T A G G C T T
1510                                     1520                                     1530

```

SUBSTITUTE SHEET (RULE 26)

## FIG.10H

GLU TYR VAL TYR ASP SER LYS GLY GLU ASN  
 GAATATGTTTATGACAGCAAGGTGAAAT  
 1540 1550 1560

LYS TRP PHE ASP ASP VAL ARG VAL SER TYR  
 AATGGTTTGAATGATGCGTGCTTAT  
 1570 1580 1590

ASP LYS GLN ASP ILE THR LEU ARG SER GLN  
 GACAAGCAAGACATTACGCTACGTAGCCAG  
 1600 1610 1620

LEU THR ASN THR HIS CYS SER THR TYR PRO  
 CTGACCAACACGCACTGTTCACCTATCCG  
 1630 1640 1650

HIS ILE ASP LYS ASN CYS THR PRO ASP VAL  
 CACATTGACAAATAATTGTACGCCCTGATGTC  
 1660 1670 1680

ASN LYS PRO PHE SER VAL LYS GLU VAL ASP  
 AATAACCTTTTTCGGTAAGAGGTGGAT  
 1690 1700 1710

ASN ASN ALA TYR LYS GLU GLN HIS ASN LEU  
 AACAAATGCCCTACAAAGAACAGCAATTTA  
 1720 1730 1740

## FIG. 10I

```

ILE  LYS  ALA  VAL  PHE  ASN  LYS  LYS  MET  ALA
A T C A A A G C C G T C T T A A C A A A A A A T G G C A
1750
LEU  GLY  ASN  THR  HIS  HIS  HIS  ILE  ASN  LEU
T T G G G C A A T A C G C A T C A C A T C A A T C T G
1770
Gln  Val  Gly  Tyr  Asp  Lys  Phe  Asn  Ser  Ser
C A A G T T G G C T A T G A T A A A T T C A A T T C A A G C
1810
LEU  SER  ARG  GLU  ASP  TYR  ARG  LEU  ALA  THR
C T T A G C C G T G A A G A T T A T C G T T T G G C A A C C
1830
His  Gln  Ser  Tyr  Gln  Lys  Leu  Asp  Tyr  Thr
C A T C A A T C T T A T C A A A A C T T G A T T A C A C C
1870
Pro  Pro  Ser  Asn  Pro  Leu  Pro  Asp  Lys  Phe
P R O P R O S E R A S N P R O L E U P R O A S P L Y S P H E
1890
Lys  Pro  Ile  Leu  Gly  Ser  Asn  Arg  Pro
A A G C C C A T T T A G G T T C A A A C A C A G A C C C
1910

```

FIG.10J

```

      ILE  CYS  LEU  ASP  ALA  TYR  GLY  TYR  GLY  HIS
      A T T G C C T T G A T G C C T T A T G G T T A T G G T C A T
      1960                                1970                                1980

ASP  HIS  PRO  GLN  ALA  CYS  ASN  ALA  LYS  ASN
G A C C A T C C A C A G G C T T G T A A C G C C A A A A C
      1990                                2000                                2010

      SER  THR  TYR  GLN  ASN  PHE  ALA  ILE  LYS  LYS
      A G C A C T T A T C A A A C T T T G C C A T C A A A A A
      2020                                2030                                2040

      GLY  ILE  GLU  GLN  TYR  ASN  GLN  THR  ASN  THR
      G G C A T A G A G C A A T A C A A C C A A T A C C
      2050                                2060                                2070

      ASP  LYS  ILE  ASP  TYR  GLN  ALA  VAL  ILE  ASP
      G A T A G A T T G A T T A T C A A G C C G T C A T T G A C
      2080                                2090                                2100

      GLN  TYR  ASP  LYS  GLN  ASN  PRO  ASN  SER  THR
      C A A T A T G A T A A C A A A C C C C A A C A G C A C C
      2110                                2120                                2130

      LEU  LYS  PRO  PHE  GLU  LYS  ILE  LYS  GLN  SER
      C T A A A C C C C T T T G A G A A A T C A A A C A A A G T
      2140                                2150                                2160

```

## FIG.10K

```

LEU  GLY  GLN  GLU  LYS  TYR  ASP  GLU  ILE  ASP
T T G G G G C A A G A A A A T A C G A C G A G A T A G A C
2170                                     2180
ARG  LEU  GLY  PHE  ASN  ALA  TYR  LYS  ASP  LEU
A G A C T G G G C T T T A A T G C T T A T A A G A T T T A
2190                                     2200
ARG  ASN  GLU  TRP  ALA  GLY  TRP  THR  ASN  ASP
C G C A A C G A A T G G G C G G G T T G G A C T A A T G A C
2230                                     2240
ASN  SER  GLN  GLN  ASN  ALA  ASN  LYS  GLY  THR
A A C A G C C A A C A A A C G C C A A T A A A G G C A C G
2250                                     2260
ASP  ASN  ILE  TYR  GLN  PRO  ASN  GLN  ALA  THR
G A T A A T A T C T A T C A G C C A A A T C A A G C A A C T
2290                                     2300
VAL  VAL  LYS  ASP  ASP  LYS  CYS  LYS  TYR  SER
G T G G T C A A A G A T G A C A A A T G T A A A T A T A G C
2310                                     2320
GLU  THR  ASN  SER  TYR  ALA  ASP  CYS  SER  THR
G A G A C C A A C A G C T A T G C T G A T T G C T C A A C C
2350                                     2360

```



## FIG.10L

THR ARG HIS ILE SER GLY ASP ASN TYR PHE  
 A C T C G C C A C A T C A G C G G T G A T A A T T A T T T C  
 2380 2390 2400

ILE ALA LEU LYS ASP ASN MET THR ILE ASN  
 A T C G C T T T A A A G A C A C A T G A C C A T C A A T  
 2410 2420 2430

LYS TYR VAL ASP LEU GLY LEU GLY ALA ARG  
 A A A T A T G T T G A T T T G G G C T G G T G C T C G C  
 2440 2450 2460

TYR ASP ARG ILE LYS HIS LYS SER ASP VAL  
 T A T G A C A G A A T C A A A C A C A A A T C T G A T G T G  
 2470 2480 2490

PRO LEU VAL ASP ASN SER ALA SER ASN GLN  
 C C T T T G G T A G A C A C A G T G C C A G C A C C A G  
 2500 2510 2520

LEU SER TRP ASN PHE GLY VAL VAL LYS  
 C T G T C T T G G A A T T T G G C G T G G T C G T C A A G  
 2530 2540 2550

PRO THR ASN TRP LEU ASP ILE ALA TYR ARG  
 C C C A C C A A T T G G C T G G A C A T C G C T T A T A G A  
 2560 2570 2580

## FIG.10M

```

SER  SER  GLN  GLY  PHE  ARG  MET  PRO  SER  PHE
AGCTCGCAAGGCTTTCGCCATGCCAAGTTT
2590
2600
SER  GLU  MET  TYR  GLY  GLU  ARG  PHE  GLY  VAL
TCTGAATGTATGGCGAACGCTTTGGCGTA
2620
2630
2640

THR  ILE  GLY  LYS  GLY  THR  GLN  HIS  GLY  CYS
ACCATCGGTAAAGGCACGCAACATGGCTGT
2650
2660
2670
LYS  GLY  LEU  TYR  TYR  ILE  CYS  GLN  GLN  THR
AAGGTCCTTTATTACATTTGTCAGCAGACT
2680
2690
2700

VAL  HIS  GLN  THR  LYS  LEU  LYS  PRO  GLU  LYS
GTCCATCAACCAAGCTAAACCTGAATAA
2710
2720
2730
SER  PHE  ASN  GLN  GLU  ILE  GLY  ALA  THR  LEU
TCCTTTAAACCAAGAAATCGGAGCGACTTA
2740
2750
2760

HIS  ASN  HIS  LEU  GLY  SER  LEU  GLU  VAL  SER
CATACCACTTAGGCAGTCTTGAGGTAGT
2770
2780
2790

```

FIG.10N

```

TYR  PHE  LYS  ASN  ARG  TYR  THR  ASP  LEU  ILE
T A T T T A A A A A T C G C T A T A C C G A T T T G A T T
2800                                2810                2820

VAL  GLY  LYS  SER  GLU  GLU  ILE  ARG  THR  LEU
G T T G G T A A A G T G A A G A G A T T A G A A C C C T A
2830                                2840

THR  GLN  GLY  ASP  ASN  ALA  GLY  LYS  GLN  ARG
A C C C A A G G T G A T A A T G C A G G C A A A C A G C G T
2850                                2860                2870                2880

GLY  LYS  GLY  ASP  LEU  GLY  PHE  HIS  ASN  GLY
G G T A A A G G T G A T T T G G G C T T C A T A A T G G G
2890                                2900                2910

GLN  ASP  ALA  ASP  LEU  THR  GLY  ILE  ASN  ILE
C A G A T G C T G A T T T G A C A G G C A T T A A C A T T
2920                                2930                2940

LEU  GLY  ARG  LEU  ASP  LEU  ASN  ALA  VAL  ASN
C T T G G C A G A C T T G A C C T A A A C G C T G T C A A T
2950                                2960                2970

SER  ARG  LEU  PRO  TYR  GLY  LEU  TYR  SER  THR
A G T C G C C T T C C C T A T G G A T T A T A C T C A A C A
2980                                2990                3000

```

## FIG. 100

```

LEU  ALA  TYR  ASN  LYS  VAL  ASP  VAL  LYS  GLY
CTGGCTTATAACAAGTTGATGTTAAAGGA
3010                                     3020
                                     LYS  THR  LEU  ASN  PRO  THR  LEU  ALA  GLY  THR
                                     AAAACCTTAACCAACTTGGCAGGAACA
                                     3040                                     3060

ASN  ILE  LEU  PHE  ASP  ALA  ILE  GLN  PRO  SER
ACATACTGTTTGATGCCCATTCAGCCATCT
3070                                     3080
                                     ARG  TYR  VAL  VAL  GLY  LEU  GLY  TYR  ASP  ALA
                                     CGTTATGTGGTGGGCTTGGCTATGATGCC
                                     3100                                     3110                                     3120

PRO  SER  GLN  LYS  TRP  GLY  ALA  ASN  ALA  ILE
CCAGCCAAATAATGGGGAGCAACGCCATA
3130                                     3140
                                     PHE  THR  HIS  SER  ASP  ALA  LYS  ASN  PRO  SER
                                     TTACCCATTCTGATGCCAAATAATCCAGC
                                     3160                                     3170                                     3180

GLU  LEU  LEU  ALA  ASP  LYS  ASN  LEU  GLY  ASN
GAGCTTTTGGCAGATAGAACTTAGGTAAT
3190                                     3200                                     3210

```

SUBSTITUTE SHEET (RULE 26)

FIG. 10P

GLY ASN ILE GLN THR LYS GLN ALA THR LYS  
GGCAACATTCACACAAACAAGCCACCAAA  
3220 3230 3240

ALA LYS SER THR PRO TRP GLN THR LEU ASP  
G C A A A T C C A C G C C G T G G C A A C A C T T G A T  
3250 3260 3270

LEU SER GLY TYR VAL ASN ILE LYS ASP ASN  
T T G T C A G G T T A T G T A A C A T A A A G A T A A T  
3280 3290 3300

PHE THR LEU ARG ALA GLY VAL TYR ASN VAL  
TTT ACC TTG CGT GCT GGCGTG TAC AAT GTA

3310                  3320                  3330

PHE ASN THR TYR THR TRP GLU ALA  
T T A A T A C C T A T A C C A C T T G G G A G G C T  
3340 3350 3360

LEU ARG GLN THR ALA GLU GLY ALA VAL ASN  
T T A C G C C A A C A G C A G A A G G C G G T C A A T  
3370 3380 3390

GLN HIS THR GLY LEU SER GLN ASP LYS HIS  
 CAGCATACAGGACTGAGCCAGATAAGCAT  
 3400 3410 3420

## FIG.10Q

TYR GLI ARG IYR ALA ALA PRO GLY ARG ASN  
 TATGGTCGCTATGCCGCTCCTGGACGCAAT 3430  
 3440  
 TYR GLN LEU ALA LEU GLU MET LYS PHE \*\*\*  
 TACCAATTGGCACTTGAAATGAAGTTTAA 3460  
 S 3470  
 3480  
 CCAGTGGCTTTGATGTGATCATGCCAATC 3490  
 3500  
 CCAATCAACCAATGAATAAGCCCCCATCT 3520  
 3530  
 ACCATGAGGGCTTTATTTTATCATCGCTGA 3550  
 3560  
 GATGCTCTTAGCCGTCTCATCTCAGATTA 3580  
 3590  
 3600  
 GTCATTAATTATAGCGATTAAATTTA 3610  
 3620  
 GTAATCACGCTGCTCTTTGATGATTTAAG 3640  
 3650  
 3660

## FIG.11A Q8 1bpB Sequence.

CCTAGGGCTGACAGTAACAACACTTTATAC 10 20 30  
AGCACATCATTTGATTTATTACCCAAATGCC 40 50 60  
ACACGCTATTATCTTTTGGGGGCAGACTTT 70 80 90  
TATGATGAATAAAGTGCCACAAGACCCATCT 100 110 120  
GACAGCTATGAGCGTCGTGGCATACGCACA 130 140 150  
GCTTGGGGCAAGAATGGGCGGGCGGCTTT 160 170 180  
TCAAGCCGTGCCCAATCAGCATCAACAAA 190 200 210  
CGCCATTACCAAGGAGCAAAACCTAACCCAGC 220 230 240  
GGTGGACAAATTCGCCAGGATAAACAGATG 250 260 270  
CAAGCGTCTTTATCGCTTTTGGCACAGAGAC 280 290 300

SUBSTITUTE SHEET (RULE 26)

## FIG.11B

ATTCACAAATGGGGCATCACGCCACGGCTG 310 320 330  
ACCATCAGCACAAACATCAATAAAGCAAT 340 350 360

GACATCAGGCAAATTATCACAAATAATCAA 370 380 390  
ATGTTTGTGTGAGTTTAGTCGCATTTTGA 400 410 420

TGGGATAAGCATGCCCCTACTTTTGTTTT 430 440 450  
GTAAAAAATGTACCATCATAGACAATATC 460 470 480

AAGAAAAATCAAGAAAAAGATTACAAAT 490 500 510  
TTAATGATAAATTGTTATTGTTTATT 520 530 540

ATTTATCAATGTAAATTGCCGTATTTGT 550 560 570  
CCATCATAAACGCATTTATCAATAATGCTCAA 580 590 600

SUBSTITUTE SHEET (RULE 26)



FIG.11C

A T A A A T A C G C C A A A T G C A C A T T G T C A A C A T 610  
 620  
 G C C A A A A T A G G C A T T A C A G A C T T T T T A G 650  
 660  
 A T A A T A C C A T C A A C C C A T C A G A G A T T A T T 670  
 680  
 690  
 MET LYS HIS ILE PRO LEU THR LEU C  
 T T A T G A A A C A C A T T C C T T T A A C C A C A C T G T 700  
 710  
 720  
 Y S VAL ALA ILE SER ALA VAL LEU LEU THR  
 G T G T G G C A A T C T C T G C C G T C T T A T T A A C C G 730  
 740  
 750  
 ALA CYS GLY GLY SER SER GLY GLY PHE ASN P  
 C T T G T G G T G G T A G C A G T G G T G G T T C A A T C 760  
 770  
 780  
 R O PRO ALA SER THR PRO ILE PRO ASN ALA  
 C A C C T G C C T C T A C G C C C A T C C C A A A T G C A G 790  
 800  
 810  
 GLY ASN SER GLY ASN ALA GLY ASN ALA GLY A  
 G T A A T C A G G T A A T G C T G G C A A T G C T G G C A 820  
 830  
 840

45/90

## FIG.11D

SN ALA GLY GLY THR GLY GLY ALA ASN SER  
 ATGCTGGCGGTACTGGCGGTGCAAACTCTG  
 850 860  
 GLY ALA GLY ASN ALA GLY GLY THR GLY GLY A  
 GTGCAGGTAAATGCTGGCGGTACTGGCGGTG  
 880 890 900  
 LA ASN SER GLY ALA GLY SER ALA SER THR  
 CAACTCTGGTGCAGGCACTGCCAGCACAC  
 910 920 930  
 PRO GLU PRO LYS TYR LYS ASP VAL PRO THR A  
 CAGAACCAAAATATATAAGATGTGCCAACCG  
 940 950 960  
 SP GLU ASN LYS LYS ALA GLU VAL SER GLY  
 ATGAAATAAAGCTGAAGTTTCAGGCA  
 970 980 990  
 ILE GLN GLU PRO ALA MET GLY TYR GLY VAL G  
 TTCAAGAACCTGCCCATGGGTATGGCGTGG  
 1000 1010 1020  
 LU LEU LYS LEU ARG ASN TRP ILE PRO GLN  
 AATAAGCTTCGTAACTGGATACCAAG  
 1030 1040 1050

SUBSTITUTE SHEET (RULE 26)

## FIG.11E

GLU GLN GLU GLU HIS ALA LYS ILE ASN THR A  
 AACAGGAAGAACATGCCAAATCAATACAA  
 1060 1070 1080

SN ASP VAL VAL LYS LEU GLU GLY ASP LEU  
 ATGATGTTGTAAACTTGAAGGTGACTTGA  
 1090 1100 1110

LYS HIS ASN PRO PHE ASP ASN SER ILE TRP G  
 AGCATATCCATTTGACCAACTCTATTGGC  
 1120 1130 1140

LN ASN ILE LYS ASN SER LYS GLU VAL GLN  
 AAACATCAAAATAGCAAGAGTACAA  
 1150 1160 1170

THR VAL TYR ASN GLN GLU LYS GLN ASN ILE G  
 CTGTTTACAACCAAGAGAGCAACACATTG  
 1180 1190 1200

LU ASP GLN ILE LYS ARG GLU ASN LYS GLN  
 AAGATCAAAATCAAAAGAGAAATAAACAA  
 1210 1220 1230

ARG PRO ASP LYS LYS LEU ASP ASP VAL ALA L  
 GCCCTGACAAAACTTGATGACGTGGCAC  
 1240 1250 1260

## FIG.11F

```

EU  GIN  ALA  TYR  ILE  GLU  LYS  VAL  LEU  ASP
TACAGCTTATATTGAAATAAGTTCTTGATG
1270                                     1280
                                     1290
                                     ASP  ARG  LEU  THR  GLU  LEU  ALA  LYS  PRO  ILE  T
                                     ACCGCTAACAGAACTTGCTAAACCCATT
1300                                     1310                                     1320

YR  GLU  LYS  ASN  ILE  ASN  TYR  SER  HIS  ASP
ATGAATAATAATTAATTAATTCACATGATA
1330                                     1340                                     1350
                                     LYS  GLN  ASN  LYS  ALA  ARG  THR  ARG  ASP  LEU  L
                                     AGCAGAAATAAGCACGCACTCGTGATTGA
1360                                     1370                                     1380

YS  TYR  VAL  ARG  SER  GLY  TYR  ILE  TYR  ARG
AGTAGTGCGTTCTCGTTATATTATCGCT
1390                                     1400                                     1410
                                     SER  GLY  TYR  SER  ASN  ILE  ILE  PRO  LYS  LYS  I
                                     CAGGTATTCTAATAATCATTCCTCAAGAAA
1420                                     1430                                     1440

LE  ALA  LYS  THR  GLY  PHE  ASP  GLY  ALA  LEU
TAGCTAAACCTGGTTTGTGATGGTGCTTAT
1450                                     1460                                     1470

```

## FIG.11G

PHE TYR GLN GLY THR GLN THR ALA LYS GLN L  
 TTTATCAAGGTACACA AACTGCTAAACAAT  
 1480 1490 1500

EU PRO VAL SER GLN VAL LYS TYR LYS GLY  
 TGGCTGTATCTCAAGTTAAGTATAAAGGCA  
 1510 1520 1530

THR TRP ASP PHE MET THR ASP ALA LYS LYS G  
 CTGGGATTTTATGACCGATGCCAAAAG  
 1540 1550 1560

LY GLN SER PHE SER SER PHE GLY THR SER  
 GACAATCATTTAGCAGTTTGGTACATCGC  
 1570 1580 1590

GLN ARG LEU ALA GLY ASP ARG TYR SER ALA M  
 ACGTCTTGCTGGTGATCGTTATAGTGCAA  
 1600 1610 1620

ET SER TYR HIS GLU TYR PRO SER LEU LEU  
 TGTCCTACCATGAATACCCATCTTTATAA  
 1630 1640 1650

THR ASP GLU LYS LYS ASN LYS PRO ASP ASN TYR A  
 CTGATGAGAAACAACAACCATATAATA  
 1660 1670 1680



FIG. 11.I

PHE THR SER ASP ALA LYS ASN SER LEU GLU G  
 T T A C C A G C G A T G C C A A A A A T A G C C T A G A A G  
 1900 1910 1920

LY GLY PHE TYR GLY PRO ASN ALA GLU GLU  
 G C G G T T T T A T G G A C C A A C G C C G A G G A G C  
 1930 1940 1950

LEU ALA GLY LYS PHE LEU THR ASN ASP ASN L  
 T G G C A G G T A A A T T C C T A A C C A A T G A C A C A  
 1960 1970 1980

YS LEU PHE GLY VAL PHE GLY ALA LYS ARG  
 A A C T C T T T G G C G T C T T T G G T G C T A A C G A G  
 1990 2000 2010

GLU SER GLU ALA LYS GLU LYS THR GLU ALA I  
 A G A G T G A G C C T A G G A A A A A C C G A G C C A  
 2020 2030 2040

LE LEU ASP ALA TYR ALA LEU GLY THR PHE  
 T C T T A G A T G C C C T A T G C A C T T G G G A C A T T T A  
 2050 2060 2070

ASN LYS PRO GLY THR THR ASN PRO ALA PHE T  
 A T A A C C T G G T A C G A C C A A T C C C G C C T T T A  
 2080 2090 2100

## FIG.11J

```

HR  ALA  ASN  SER  LYS  LYS  GLU  LEU  ASP  ASN
   CCGCTAACAGCAAAAGAACTGGATAACT
   2110
      PHE  GLY  ASN  ALA  LYS  LYS  LEU  VAL  LEU  GLY  S
      TTGGCAATGCCCAAAGTGTGCTCTTGGGTT
      2140
ER  THR  VAL  ILE  ASP  LEU  VAL  PRO  THR  GLY
   CTACCGTCATTGATTGTGTCCTACCGGTG
   2170
      ALA  THR  LYS  ASP  VAL  ASN  GLU  PHE  LYS  GLU  L
      CCACCAAGAATGTCAATGAATTCAAGAAA
      2200
YS  PRO  LYS  SER  ALA  THR  ASN  LYS  ALA  GLY
   AGCCAAAGTCTGCCACAACAAGCGGCG
   2230
      GLU  THR  LEU  MET  VAL  ASN  ASP  GLU  VAL  ILE  V
      AGACTTTGATGTTGAATGATGAGTTATCG
      2260
AL  LYS  THR  TYR  GLY  TYR  GLY  ARG  ASN  PHE
   TCAAAACCTATGGCTATGGCAGAACTTTG
   2290
      2300

```



## FIG.11K

GLU TYR LEU LYS PHE GLY GLU LEU SER ILE G  
 AATACCTAAATAATTGGTGAGCTTAGTATCG  
 2320 2330 2340

LY GLY SER HIS SER VAL PHE LEU GLN GLY  
 GTGGTAGCCATAGCGTCTTTTACAAGCGG  
 2350 2360 2370

GLU ARG THR ALA GLU LYS ALA VAL PRO THR G  
 AAGCACCGCTGAGAAAGCCGTACCAACCG  
 2380 2390 2400

LU GLY THR ALA LYS TYR LEU GLY ASN TRP  
 AAGCACAGCCAAATATCTGGGGAAC TGGG  
 2410 2420 2430

VAL GLY TYR ILE THR GLY LYS ASP THR GLY T  
 TAGGATACATCACAGGAAGGACACAGGAA  
 2440 2450 2460

HR SER THR GLY LYS SER PHE ASN GLU ALA  
 CGAGCACAGGAAAGAGCTTTAATGAGGCC  
 2470 2480 2490

GLN ASP ILE ALA ASP PHE ASP ILE ASP PHE G  
 AAGATATTGCTGATTTTGACATTGACTTTG  
 2500 2510 2520

## FIG.11L

```

LU  ARG  LYS  SER  VAL  LYS  GLY  LYS  LEU  THR
AGAGAAATCAGTTAAAGGCCAAACTGACCA      2530
                                     2540
                                     2550
                                     THR  GLN  GLY  ARG  GLN  ASP  PRO  VAL  PHE  ASN  I
                                     CCCAAGGCCGCCAAGACCCCTGTATTTAACA      2560
                                     2570
                                     2580

LE  THR  GLY  GLN  ILE  ALA  GLY  ASN  GLY  TRP
TCACAGGTCAAAATCGCAGGTAAATGGCTGGA      2590
                                     2600
                                     2610
                                     THR  GLY  THR  ALA  SER  THR  ALA  LYS  ALA  ASN  V
                                     CAGGCACAGCCAGCACCCGCCAAGCGAACG      2620
                                     2630
                                     2640

AL  GLY  GLY  TYR  LYS  ILE  ASP  SER  SER  SER
TAGGGGGCTACAGAATAGATTCTAGCAGTA      2650
                                     2660
                                     2670
                                     THR  GLY  LYS  SER  ILE  VAL  ILE  GLU  ASN  ALA  L
                                     CAGGCAAAATCCATCGTCAATCGAAATGCCA      2680
                                     2690
                                     2700

YS  VAL  THR  GLY  GLY  PHE  TYR  GLY  PRO  ASN
AGGTACAGGTGGCTTTTATGCTCCAAATG      2710
                                     2720
                                     2730

```

SUBSTITUTE SHEET (RULE 26)

## FIG.11M

ALA ASN GLU MET GLY GLY THR HIS A  
 C A A C G A G A T G G G C G G G T C A T T T A C A C A C G 2760  
 2740 2750

SP THR ASP ASP SER LYS ALA SER VAL VAL  
 A T A C C G A T G A C A G T A A A G C C T C T G T G G T C T 2790  
 2770 2780

PHE GLY THR LYS ARG GLN GLU VAL LYS \*  
 T T G G C A C A A A A G A C A A G A G A G T T A A G T 2820  
 2800 2810

\*\*

A G T A A T T T A A C A C A A T G C T T G G T T C G G C T 2850  
 2830 2840

G A T G G G A T T G A C G C T T A A T C A A A C A T G A A T 2880  
 2860 2870

G A T T A A G A T G A T A A A C C C A A G C C A T G C C A A 2910  
 2890 2900

T G A T T G A T A G C A A C G A T G G C A G A T G A T G A G 2940  
 2920 2930

T T T T C A T T A T C T G C C A T T A T T A T T G C T T A A 2970  
 2950 2960

T T A T T G C T T G T C A T T T G G T G G T G T A T C A C 3000  
 2980 2990

## FIG.11N

A T T A A T C A T : A A A A T T A A C A T A A T A A T G A 3010 3020 3030  
 T T A A A T G A T A T T T A A T G A A A G T C A G G G T T A 3040 3050 3060  
 T T T T G G T C A T G G T T T T C A T G A T T A T T T A A 3070 3080 3090  
 C T T A T A A T G C G T T A T G G T T A G C A A A A G C T 3100 3110 3120  
 A A G T C T G T C A A T G A A G C T A T G G T G A G T G A T 3130 3140 3150  
 T G T G C A A A A G A T G G T C A A A A A A T C G G T A T 3160 3170 3180  
 G G T G C T G T C A G G C G T G G T G A T G G T T C T G T T 3190 3200 3210  
 A A T G A T A A T A A C A A C G C C A A G C C A T G C T A C 3220 3230 3240  
 T G C C A A G T T G T T G C C G A C C T C T C A A G A A A A 3250 3260 3270  
 T C C A A C C A A A A C T A T G G T A G A T A G C T T T G G 3280 3290 3300

## FIG. 11.O

TCGTGAACGCCACGAGGGGCGAGTTCAGGG 3310 3320 3330  
GCTATTGCCGTGCAATTGCAGCAGAAAGACTA 3340 3350 3360  
TGAGCTGGCTGCCAACTATTGGACGGCCG 3370 3380 3390  
TTATTGGCAAAACCCAAACGCCCAATCG 3400 3410 3420  
TGAGATTGTTGAGCA 3430

SUBSTITUTE SHEET (RULE 26)

FIG. 12A

Tbop1 alignment

	10	20	30	40	50	60	
MNOSKONNKSKQVILKSALSIGLINI---	TQVALANTTADKAEA-TDKTNLWVLDETIVVT						
.Q.QHLFR-----	NILC....	MT.PVY----	NVQAEOAQEQ..TIQ.K				
.Q.QHLFR-----	NILC....	MT.PAY----	NVQAEOAQEQ..TIQ.K				
.Q.QHLFR-----	NILC....	MT.PAY----	NVQAEOAQEQ..TIQ.K				
TKPKPYFR-----	LSIIISC.LI.CYVKAE..SIKDITKE.ISS.VD.QS.E-DSE.ETIS..						
		70	80	90	100		
							AKKVA-RKANEVITGLGVVKTAEITINKEQVINRDLTRYDP
							.....
							...QKT.RD.....L.SSD.LS.....
							...QKT.RD.....L...D.LS....D.....
							...QKT.RD.....L...D.LS....D.....
							E.IRD..D.....II..S.S.SR.....
							Eagan

	110	120	130	140	150	160
GLAWEQCFGASSYSIRGMCDKNRVAVLDGINQAQHVALQPVAGKNYA-AGGAINEIEYEN ..... SLT....VS.I.S.TA.AALG.TRT.GSS. ..... SLT....LA.I.S.TA.AALG.TRT.GSS. ..... SLT....LA.I.S.TA.AALG.TRT.GSS. S.....R.....LP.T.S.W.S.LVATSGYST.						

FIG.12B

170 180 190 200  
 VRSVEISKANSSEYCGSALSGSVAFVTKTADDIKDG  
 .....  
 .KA.....S.....N...A.....Q....A...GE.  
 .KA.....S..V.Q.....A.....Q.....V.GE.  
 .KA.....S..V.Q.....A.....Q.....V.GE.  
 .KA.....GS.....N...A...T.QS.S.A..LEGD

4223  
 Q8  
 B16B6  
 M982  
 FA19  
 Eagan

210 220 230 240 250 260  
 KDWGVQTKTAYASKNNWNSVAAGKAGSFGSLIYTDRCQEQYKAHDDAYQGSQSFDRVA  
 .....  
 .Q..I.S....SG.DH.LTQ.L.L..RS.GAEA.L...K...R.IH..K..GK.V...N.L.L  
 RQ..I.S....SG..RGLTQ.I.L..RI.GAEA.L.H.G..AG.IR..E..GR.V...N.L.P  
 RQ..I.S....SG..RGLTQ.I.L..RI.GAEA.L.H.G..HAG.IR..EA.GR.V...N.LAP  
 .S...I...N.S....KGFTH.L.V...Q.G.E..A...Q.NSI.TQV.K..LK.V..Y..LI.

270 280 290 300  
 TTD-----FNNRTFLIANECANFNYEACAAGQTKLQAKPTN  
 .....PK.....  
 DE.KREGGSQY.Y.IVEE..H..-..A..KNKL--..ED.SVKD  
 VE.-----SSEYAY.IVED..EGK...T.KSKP--..KDVVGKD  
 VE.-----GSKYAY.IVEE..K..GH.K.K.NP--..KDVVGED  
 ....-----KSSGY.V.QG..P..DDK-....-PP.TLST

4223  
 Q8  
 B16B6  
 M982  
 FA19  
 Eagan

FIG.12C

310	320	330	340	350	360	
VRDKVNVKDYTGPNRLIPNPLTQDSKSLILRPGYQLNDK-HYVGGVVEITKQNYAMQDKTVPA						
.....						
E.KT.STQ.....S...LA...EYG.Q.W.F...WH.DNR-.....A.L.R.Q.TFDIR.M....						
E.QT.STR.....FLAD..SYE.R.W.F...FRFENKR..I...IL.H.Q.TFDIR.M....						
K.QT.STR.....FLAD..SYE.R.W.F...FRFENKR..I...IL.R.Q.TFDIR.M....						
QSET.S.S....A.IK...MKYE.Q.WF..G..HFSEQ-.I...IF.F.Q.KFDIR.M.F..						
	370	380	390	400		
	YLTVHDIEKSRLSNHAQA--NGYYQCNILGERIRDITIGPD					4223
	.....G.---					Q8
	.F.SE.YVPGS.KGL-----K.S.D.KA..LFVQGECS					B16B6
	F..KAVFDANSKQAGSLPG-.K.A..HKYGGGLFTINGENG					M982
	F..KAVFDANQKQAGSLPG-.K.A..HKYGGGLFTSGENN					FA19
	..SPTERDDSSRSFYPMQDH.A..HIE-----					Eagan
410	420	430	440	450	460	
-----SGYGINYAHGVFYDEKHQKDELGLEVYVDSKGENKWFDDVRVSYDKQDITILRSQJINTHC						
-----						
TLQGI.....T.....R.T.N.Y.V...HNADKDT.A.YA.L...R.G.D.DNR.QQ....						
-----ALV.AE.GT.....T.T.S.Y...TNADKDT.A.YA.L...R.G.G.DNHFFQQ....						
-----APV.AE.GT.....T.T.S.Y...TNADKDT.A.YA.L...R.G.G.DNHFFQQ....						
-----D.R.VK..S.LYF..H.R.Q.V.I..I.EN.NKAGII.KAVL.ANQ.N.I.D.YMRH....						



FIG.12D

470	480	490	500	
STYPHIDKNCCTPDV	NKPF	SVKEVDNNAYKEQ	NLIKAVFN	4223
.....	.....	.....	.....	Q8
.HDGS-.....	R..G..Y.FYKS.RMI.E.SR..FQ...K			B16B6
.ADGS-.....	Y.R.SAD....YKKS.RVI.G.S.R.LQ.A.K			M982
.ADGS-.....	Y.R.SAD....YKKS.RVI.G.S.K.LQ.A.K			FA19
.L..NPS.....	R.TLD..Y.YYRS.R.V...K..MLQLNLE			Eagan

510	520	530	540	550	560
KKMALGSTHHHINLQ	GYDKFNSSL	SREDVRLATHQ	SYQKLDYTPPSNPLDKF-KPILGSNN		
.....N.....	.....	.....	.....	.....	.....
.AFDTAKIR.NLSINL...	R.K.Q..HS..Y.QNAVQAYD.I-...	KP.F.NGS-..-----D			
.SFDTAJUR.NKSVNK.F.R.S.B.RHQ..	YYQHANRAYSSK-...	KTAN.NGD-..-----S			
.SFDTAKIR.NLSVNL...	T.G.N.RHQ..YYQSANRAYS.K-...	Q.NGKTS-..--FN.REK			
..IQQWMLT.Q.VFNL.F.D.T.A.QHK..	-..TRRVIAATA-.SI.RK---	GETG..RN.LQS			

570	580	590	600	
KPICLDAYGYGHDHPQACNAKNSTYQNFAIKGIEQYN				4223
R.....	.....	.....	.....	Q8
N.YRVSIGK-----				B16B6
..YWSIG-----				M982
N.YWSIGR-----				FA19
Q.YLYPKPEP-----				Eagan

FIG.12E

610	620	630	640	650	660
QKTNIDKIDYQALIDQYDKQNPINSTLKPFEEKIKQSLGQEKYNKIDELGFKAYKDLRNEWAGWT					
.....V.....					
-----DE..R..N.....					
-----					
-----					
-----					

670	680	690	700
NDNSQQVANKGRDNIYQPNQA-TVVKDDCKKYSETNS-Y			
.....			
-----T..NTSPI.RFCN-.T-. .			
-----GN..TGQI.LFCN-.T-. .			
-----GN..TRQI.LFCN-.T-. .			
-----YFAGQDH-.N.QGSS.N.			

4223  
Q8  
B16B6  
M982  
FA19  
Eagan

710	720	730	740	750	760
ADCSTTRHISGDNFYALKONMTINKYVDLGLGARYDRIKHSIDVPLVDNSASNQLSMNFGVW					
.....					
T...-P.N.G.NG.YA.VQ..VRLGRWA.V.A.I...YRSTH.EDKS.STGTHRN....A....					
T...-P.S.N.KS.YA.VR..VRLGRWA.V.A.L...YRSTH..DCS.STGTHRT....A.I..					
T...-P.S.N.KS.YA.VR..VRLGRWA.V.A.L...YRSTH..DCS.STGTHRT....A.I..					
R...-KV.L.K.K..YF.ARN..ALG.....I...VSRT.ANESTISVGKFNF....T.I..					

FIG.12F

770	780	790	800			
VKPTNWL	DIAYRSSQCFRMP	SFSEMYGERFGV	TIGKG	4223		
.....	.....	.....	.....	Q8		
L..FT.M.LT.	A.T...L...A....	W.A.ESL	KTL	B16B6		
L..AD...LT.	T.T...L...A....	W.S.OSK	AV	M982		
L..AD...LT.	T.T...L...A....	W.S.DK.	KAV	FA19		
I...E....LS.	L.T...N.....	W.Y.GKND	EV	Eagan		
810	820	830	840	850	860	
TOHGCKGLY	ICQIVHQTKLPEK	SFNOEIGATLHNL	GSLEVSFKNRY	TDLIVKSEEIR		
.....	.....	.....	.....	.....	.....	
-----D.....	R.A.IVFKGDF.N.	A...N.A.R....	AFGY-.T.			
-----ID.....	K.A.IVFKGDF.N.	A.W.N.A.R....	RGY.AQI			
-----ID.....	K.A.IVFKGDF.N.	A.W.N.A.R....	RGY.AQI			
-----YVG.F...	T.R...F.LA.KGDF.	NI.I.H.S.A.RN.	AFA-.LS			
870	880	890	900			
TLTQGNAGK	QKRGKDLGFHNGQ	DADLTGINILGR	LD	4223		
.....	.....	N.....	K..	Q8		
-----QN.QTSAS.	P.YR.A.N.RIA....	KI.		B16B6		
K-----N..EEA...	PAYL.A.S.RI....	KI.		M982		
K-----D..EQV..	NPAYL.A.S.RI....	KI.		FA19		
K-----NGT...	NY.Y..A.N.K.V.V..	TAQ..		Eagan		

FIG. 12G

910	920	930	940	950	960
INAVNSRLPYGLYSTLAYNKVDVKGKTLNPTLAG-TNILEDAIQPSRYVWGLGYDAPSQKWGA					
.....					
WHG.WCG.D.....	RIK..	DADIRADRTFV.SY.....	V.....	L.....	H.DGI..I
WNG.WDK..E.W...F....	R.H.RDIKKRADRTDIQSH.....				Q.EG..V
WNG.WDK..E.W...F....	R.H.RDIKKRADRTDIQSH.....			S.....	Q.EG..V
F.GLMK.I...W.A.F..Q.K.DQKI.AG..SVSSY.....				II.....	H.NT..I

[illegible]

FIG. 13A

## Top2 comparison

	10	20	30	40	50	60
MKHIPLTILCV	SAISAV-LLTACGS-GGSNPPAPTPI	PNVAGSGN	ICN	ICN	AGG	TNT-ANAG
.....	.....	S.F...	S.....	GN...A.A.....	GCANSG...	
NN-	VNQAA	M	L	P.F.S.L.G-		
NN-	VNQAA	M	L	P.F.S.L.G-		
NN-	VNQAA	M	L	P.F.S.L.G-		
SV	ISGELS	--F.S.S.				

70	80	90	100
NIGGT---NSGTGSANTPEKYQDVPTKNEKD-K-VSSIQEPAM			
.A...GGA...A.S.....K....DE.K.AE-.G.....			
-FDLDSVE---VQDMHSK...EDEKS-QP.SQOD..ENSGA.-			
-FDLDSVD---EAPRPA-.....SS.PQAQ.D-----QG			
-FDLDSVD---EAPRPA-.....PSK.P.AR.D-----QG			
-FDVDNW---N.P.---SK.R...DTSNORK.S-NLKLF.I.SL			

110	120	130	140	150
GYCMALSKINLHNRQDTPLD-EKNIITL--DGKKQVAEG-KKSPLPFS-LDV-ENKLLDGYIA				
VE.-.LRMWIP.EQEETH-A.IN-.N--.VV.LEGDL-.HN.FDN.IWQNIK.SKEVQTVY				
-...F.V-.LPRR.AHFN.KYK..HKP.GSM.W-----.-LQGEPSNFS.RDE.E-----				
...F.M-RLKRR.WYP--GAE.SEVK.NES.WEATGLPTKP.E-..KRQKS.I.KVET..D-S				
...F.M-RFKRR.WHPSANPK.DEVK.KND.WEATGLPTEP.K-..LKQQS.ISEVEIN.N-S				
.G.K.VAQ..RCNKEPSFLN.DDY.-----SY..S.STI.KDVK.NNK-				

FIG.13B

160	170	180	190	200	
KNVADKNAIGDRIKGNKEISDEELAKQIKEAVRKSHEFQQV-					4223
NQEKQNIEDQIK.EN.QRPDKKLLDDV.L.AYIEKVLDDRLTELA					Q8
-----K.R.SS.LI-.SKWEDGQSR.VGYTN.T----					B16B6
DIYSSPYLTPSNHQNG-----AGNGVN.P.NQATGHEN.---					M982
..YTSPYLSQDADS-----HANG.N.P.NE.TDYKK.-----					FA19
---G.--L..S-.-----..PSTINPP.K-----HG.---					Eagan

210	220	230	240	250	260
LSSLENKIFHSNDGTTKATTRDLKYVDYGY-VLANDGNVLTVKTDKLNILGPFVGGVFNCTTT					
KPIY.KN.NY.H.KQN..R.....RS..I.RSGYS.---IIPK.IAKT.FD.AL..Q..Q.					
-----RS..V..-..KN.IDIKNNIV.F--..D.VLY.K.KEP					
-----..YS.WF.KH.ASEKDFSN.KI.S---.DD.YI..H.EK					
-----..YS.WF.KH.KSEVKNEGLVSAKR-..D.YI..H.DK					
-----..YS.LY.TPSWSLND-S..N-.FY-..YY.YA..Y.NK.					

270	280	290	300	
AKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENS--QA				4223
..Q..VSQ-.....T.....-..KKGQS..SFCT-.QRL.				Q8
S.....SEKIT...T..YV..AME-KQ..-..GLG-.A..G				B16B6
PSRQ..ASGK.I...V.H.V..TKGQD.R.IIQP.KK.G				M982
PSRQ..ASE..T...V.H.V..TKQGQK.NDIL.T.KG.G				FA19
.TN..VNGVA....T...I.ATK.-GK.YPLLSNG.H.---				Eagan

FIG.13C

310	320	330	340	350	360	
GWYYGASSKD-EYNRLITKEDSAPDGHSGEYGHSSSEFTVNFKEKLTGKLFN---						
.DR.S.M.YH-.PS...D.KNK..NVN.....D.SK.S.K.E.S.---I.G...S						
DK-S..L.AL-.EGV.RNQAE-ASS..TD-F.MT...E.D.SD.TIK.T.YR.NRIT.NNSENK						
DR.S.F.GDGS.EYSNKN-.STLK.D.E.-..FT.NLE.D.GN.....IR.NAS.NNNINND						
DK.S.F.GDEG.TTSNR.-DSNIN.K.E.-..FT.N.K.D.NN.....IR.NKVINPAASDG						
--.RR-.AIP.DID.EN-DSKNG.-I.-----LI...SADGGT.....Q.-.YTKRKTNNQPYE						
370	380	390	400			

VTKTERYDIDANIHGVRFRGSATASNK--NDTSK-HPFTSDAN						4223
.N.K.....Y.....DITEASK...-.....K						Q8
QI..T..T.Q.TL....K.K.L.AD.--GA.NGS...I...SD						B16B6
KHT.QY.SL..Q.T...N.T...TD.K-ENET.L...V...SS						M982
Y...Y.SL...TLR....S.K.I.TD.PNTGGT.L...VF.SS						FA19
KK.L--.....D.YS.....TVKPTE.---SEE-.....EGT						Eagan

410	420	430	440
NRLEGGFYCPKGEELAGKFLTNDAKLFQVFGAKRESK-----AEKTE-----			
.S.....NA.....E-----K.....			
S-.....S.....VAA.....QKD.KDGENA.GPA.,-----			
S-.S...F...Q....GFR..SD.Q.VAV.GS..TKD.LENGAA.SGS.G-AAASGGAAGTSSE			
S-.S...F...Q....GFR..SD.G.VAV.GS..TKDST-----NGNAP-AASSGPGAATMPS			
--.....NA...G....AT..RV....S..ETEETKKEALSK.TLIDGKLITFSTKKTDA			

**SUBSTITUTE SHEET (RULE 26)**

4223  
Q8  
B16B6  
M982  
FA19  
Eagan



**SUBSTITUTE SHEET (RULE 26)**

4223  
Q8  
B16B6  
M982  
FA19  
Eagan

FIG.13F

680	690	4223
PVANEMGGSFT-----	NADDSKASV	Q8
.....	HDT.....	B16B6
K..I.....	SFPCNAPECKQE-----	M982
.K.E.L..W.AYPGDKQTEKATATSSDG---	SAS.-.T.	FA19
...E.L..W.AYPGNEQTKNAIVESGNG---	SAS.-.T.	Eagan
.K.S.L..Y..YNGNSTATNSESSTVSSSS.SKNAP.A.		

700	4223
VFGTRRQQEV-K*	Q8
.....E...*	B16B6
...A....L.Q-*	M982
...A....P.Q-*	FA19
...A....KL.-*	Eagan
...ARQ.V.TT.*	

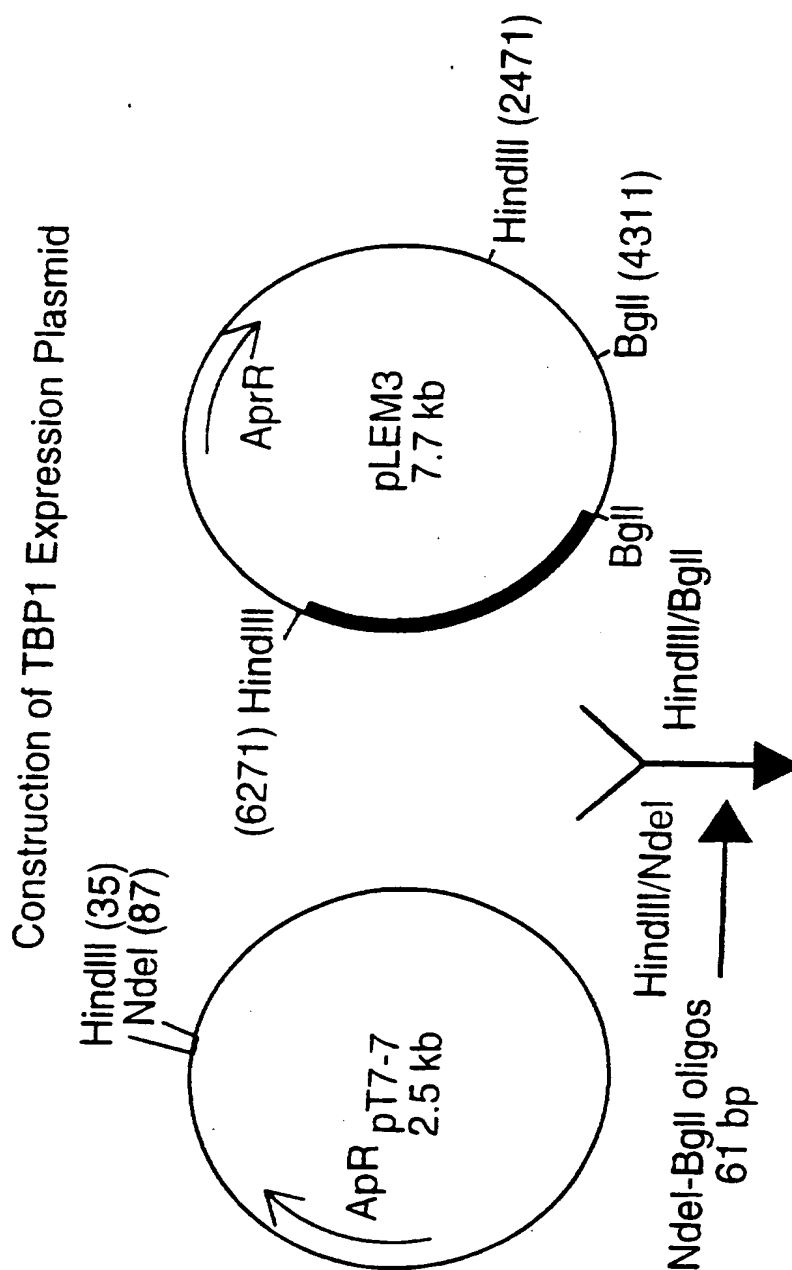


FIG.14A

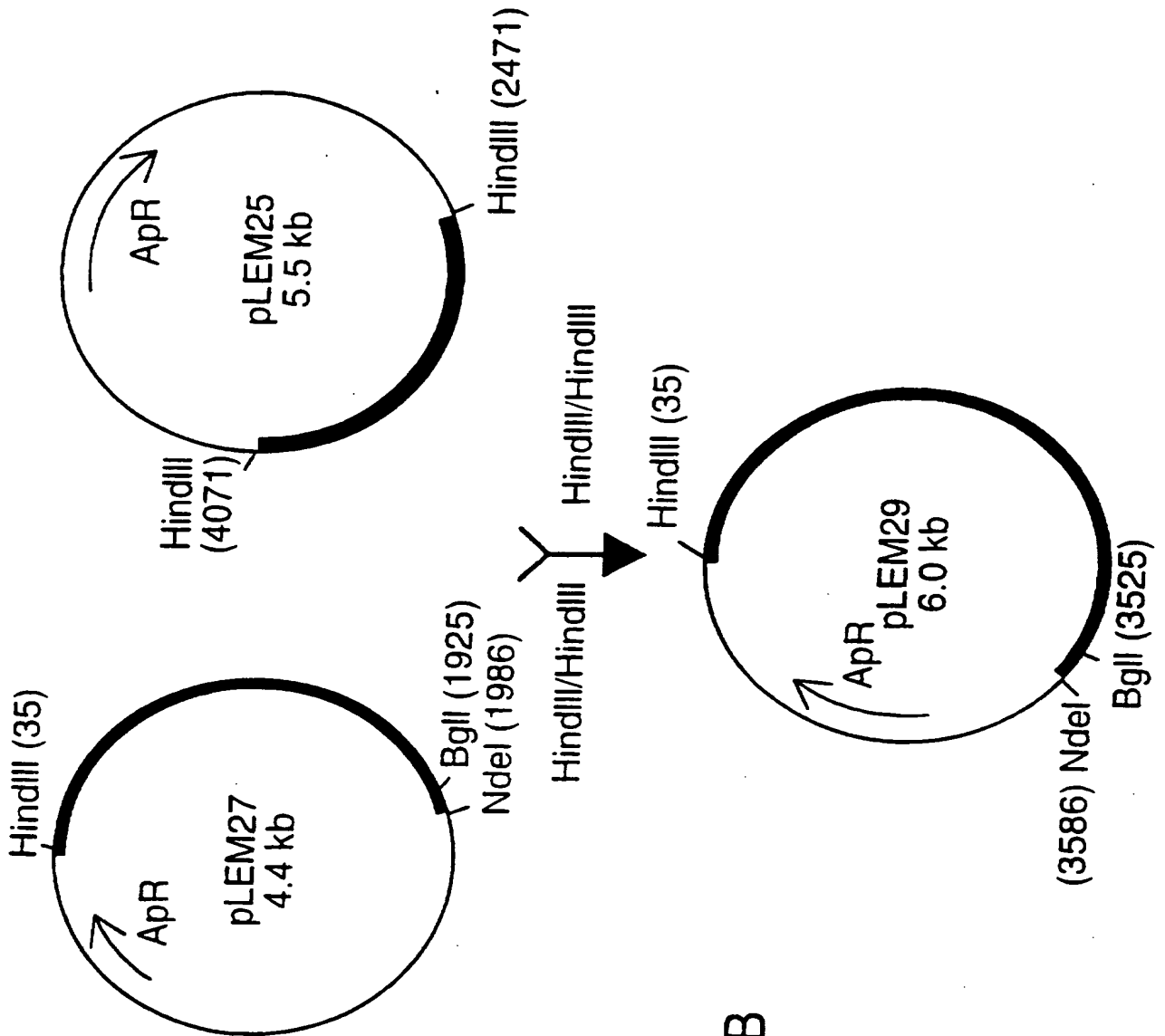
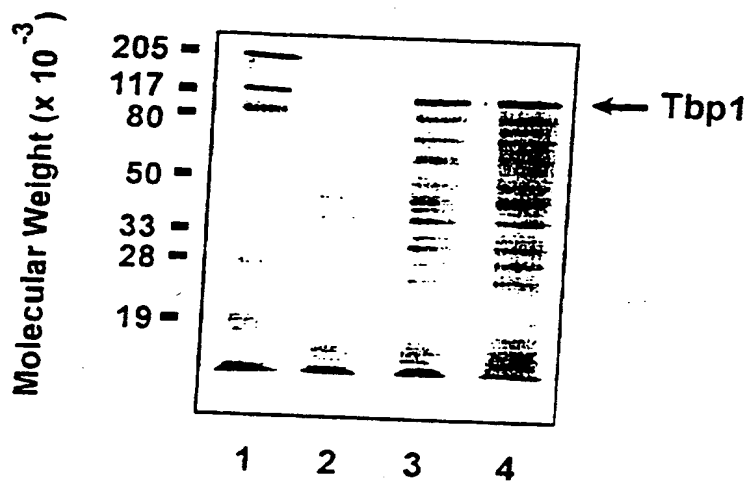


FIG.14B

## Expression of rTbp1 in *E. coli*



1. Prestained molecular weight markers
2. pLEM29B-1 lysate, non-induced
3. pLEM29B-1 lysate, 1 hr post-induction
4. pLEM29B-1 lysate, 3 hr post-induction

Fig.15

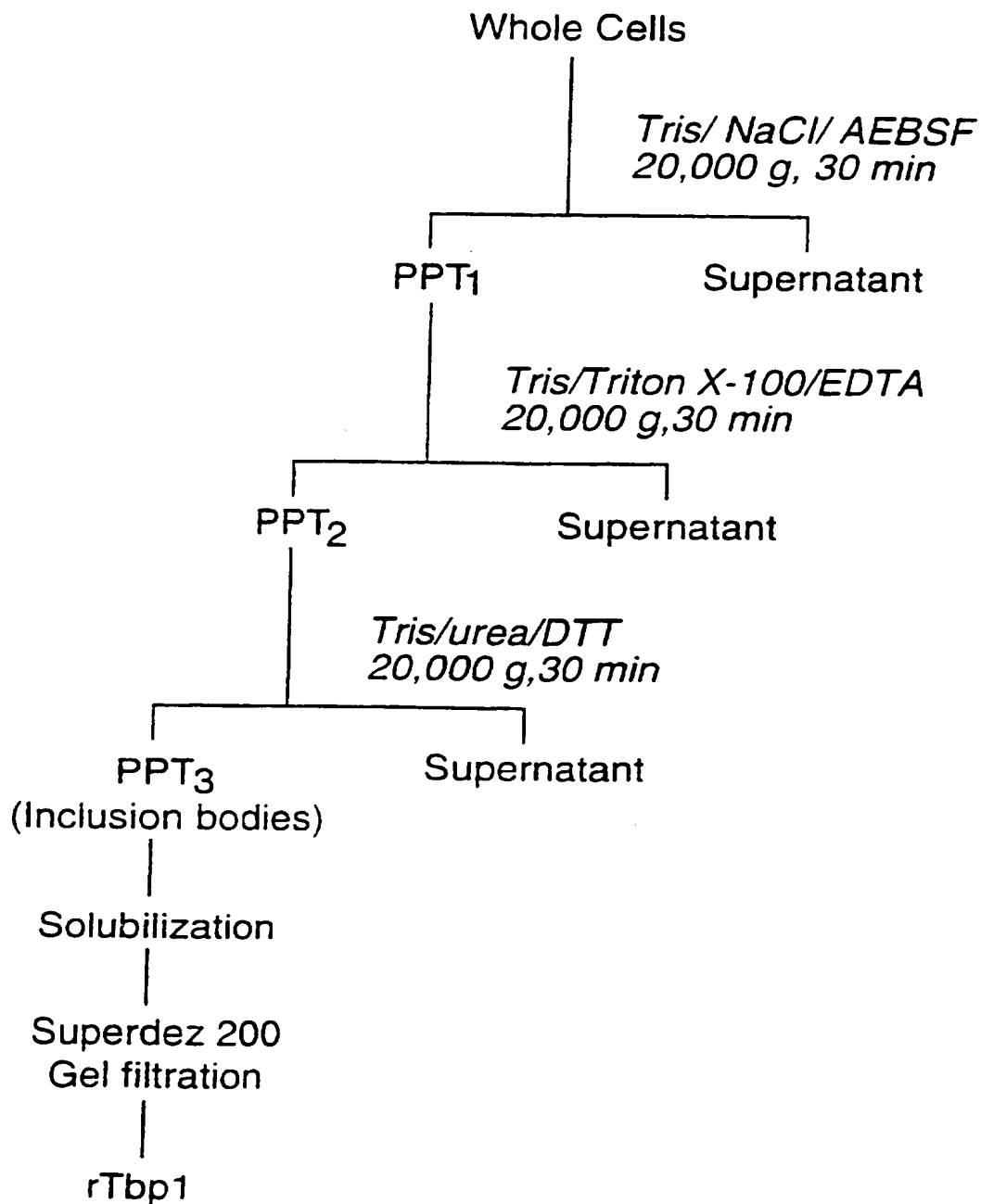
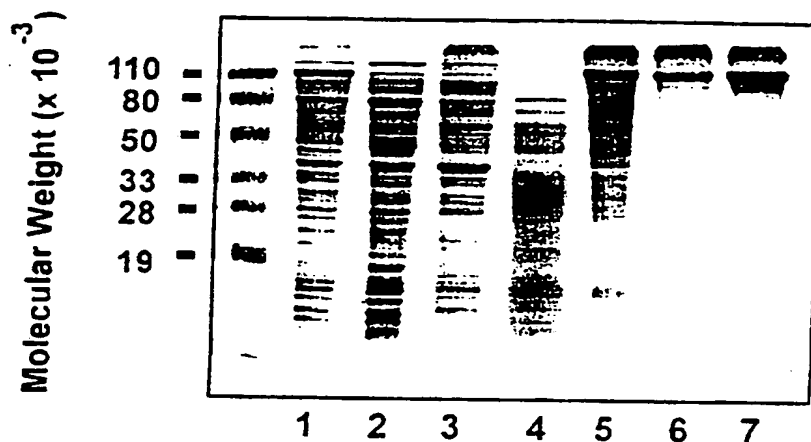
Purification of Tbp1 from *E. Cole*

FIG.16

## Purification of rTbp1 from *E. coli*



1. *E. coli* Whole cells
2. Soluble proteins after 50 mM Tris/ NaCl extraction
3. Soluble proteins after Tris/ Triton X-100/ EDTA extraction
4. Soluble proteins after Tris/ urea/ DTT extraction
5. Left-over pellet (rTbp1 inclusion bodies)
- 6.7. Purified rTbp1

Fig.17

CONSTRUCTION OF TBP2 EXPRESSION PLASMID

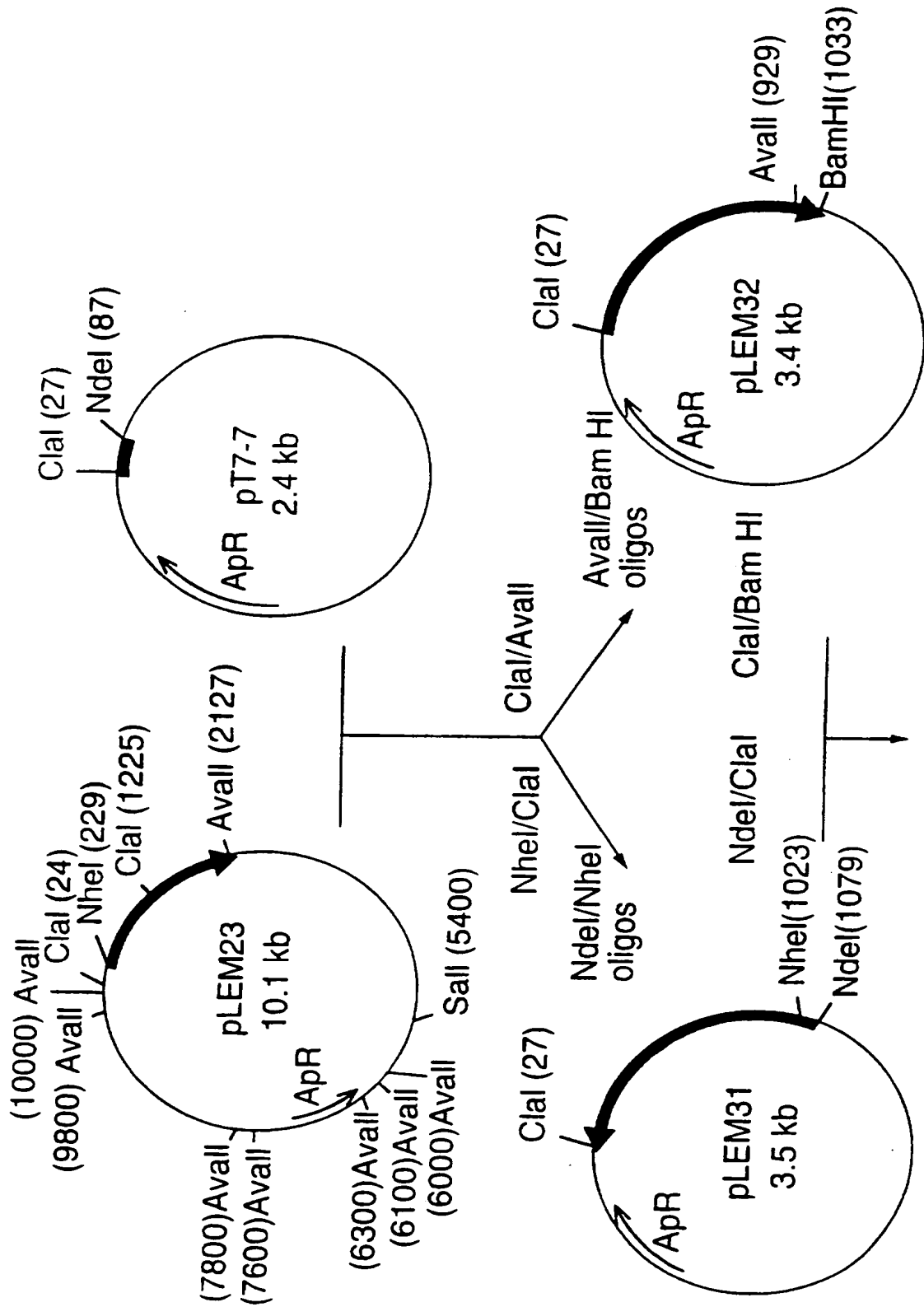


FIG.18A



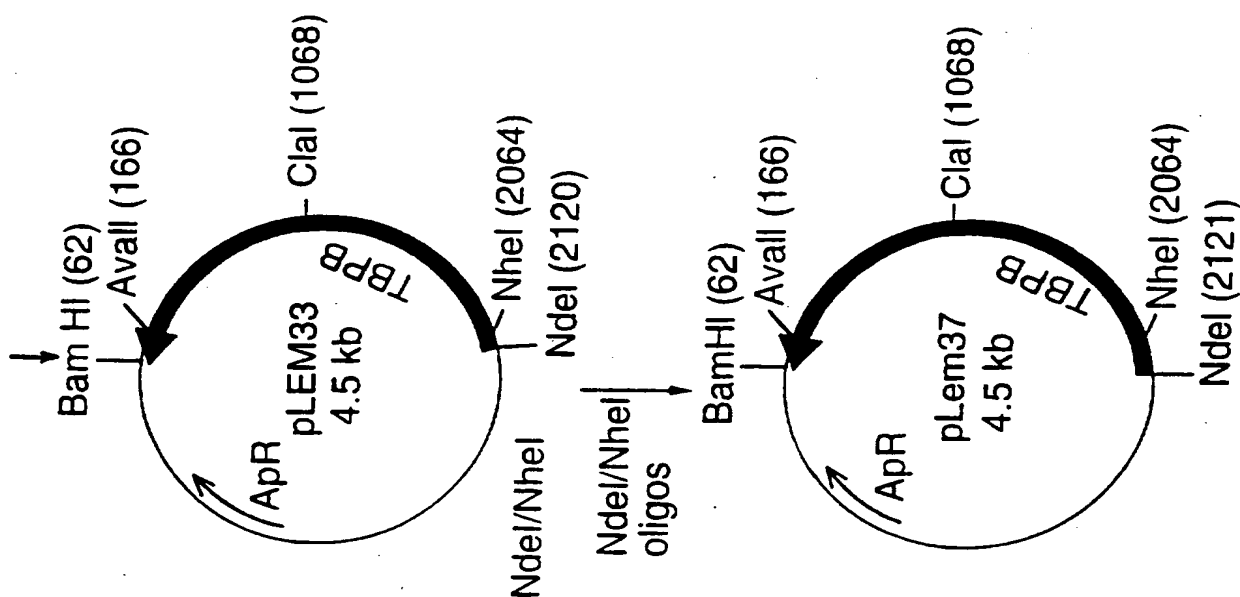
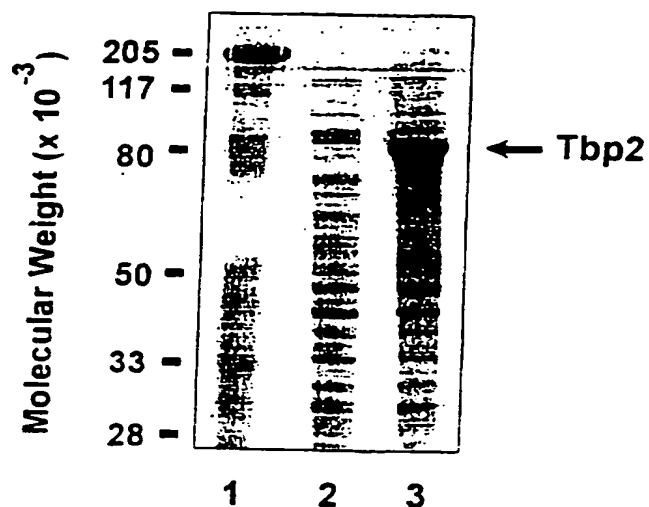


FIG.18B

## Expression of rTbp2 in *E. coli*



1. Prestained molecular weight markers
2. pLEM37B-2 lysate, non-induced
3. pLEM37B-2 lysate, induced

Fig.19

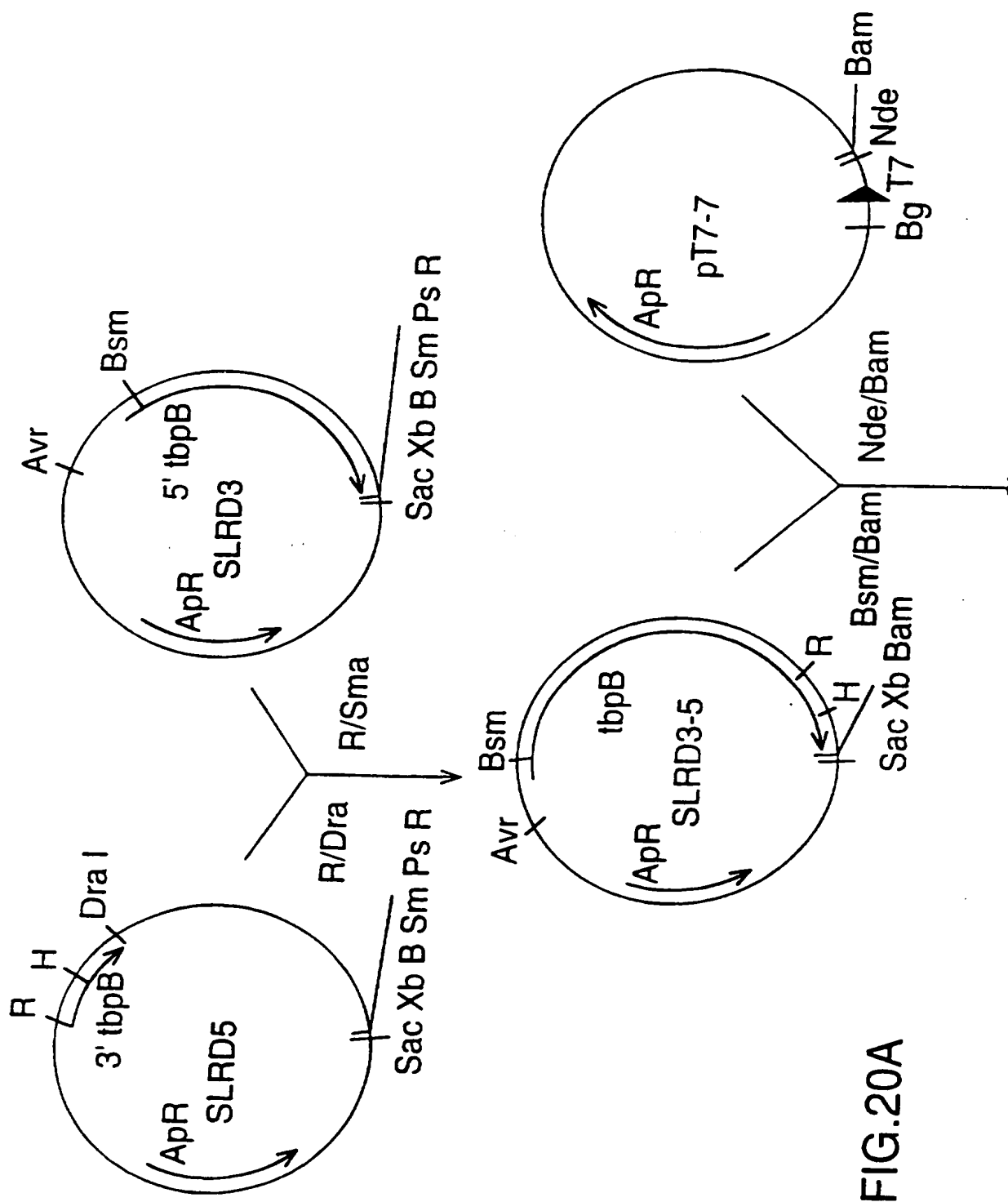


FIG. 20A

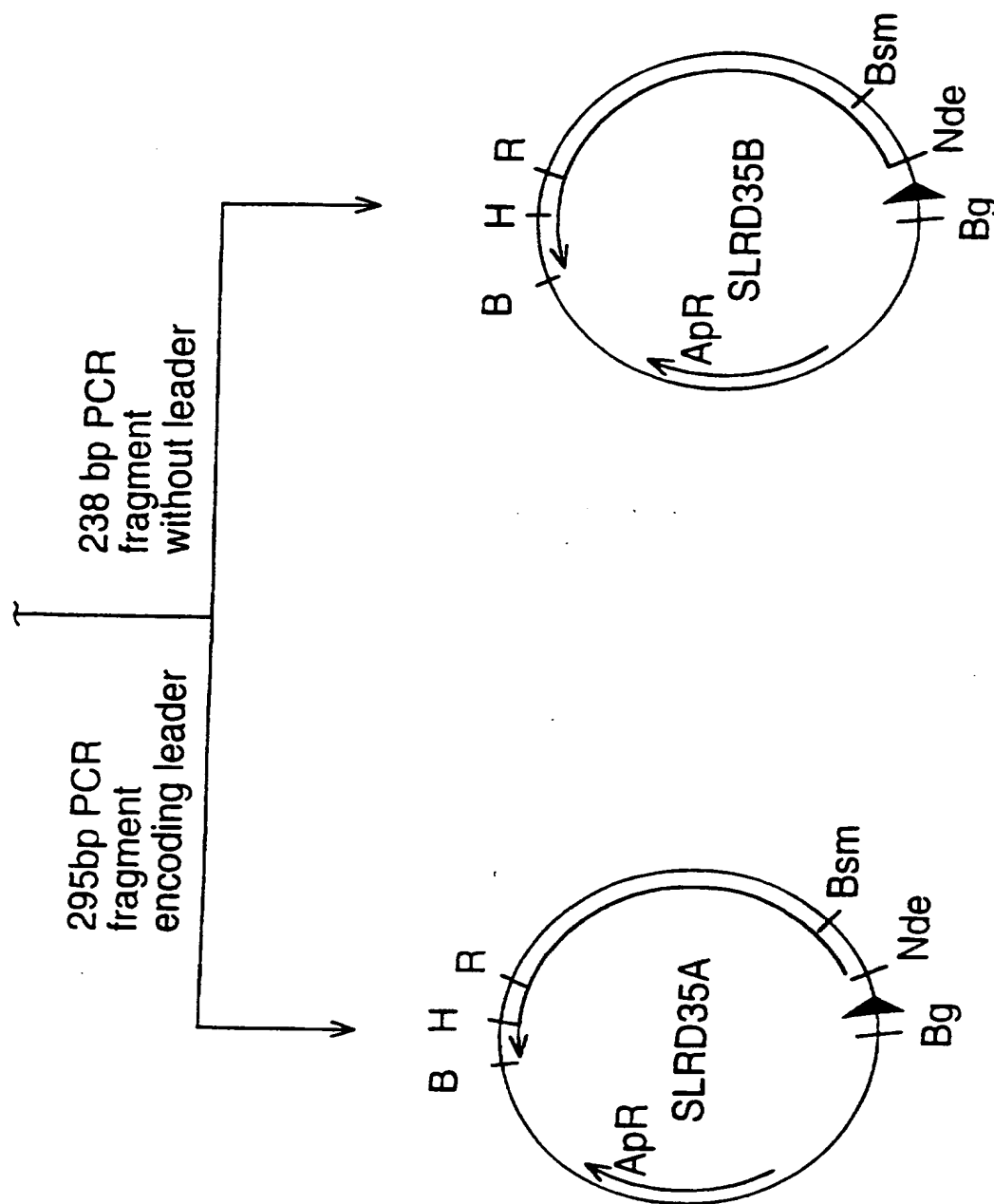
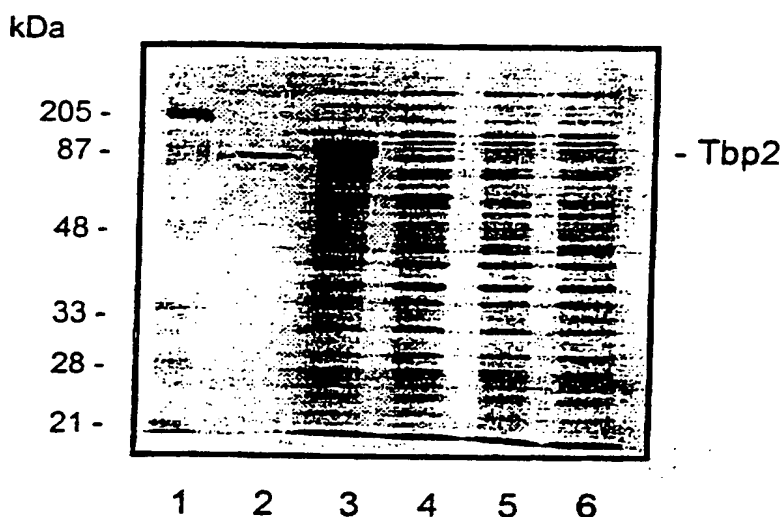


FIG.20B

**Fig 21. Expression of Q8 rTbp2 protein in *E. coli***



1. Prestained molecular weight markers
2. 4223 rTbp2 protein
3. SLRD35A lysate, 3 hr post-induction
4. SLRD35B lysate, 3 hr post-induction
5. SLRD35A lysate, non-induced
6. SLRD35B lysate, non-induced

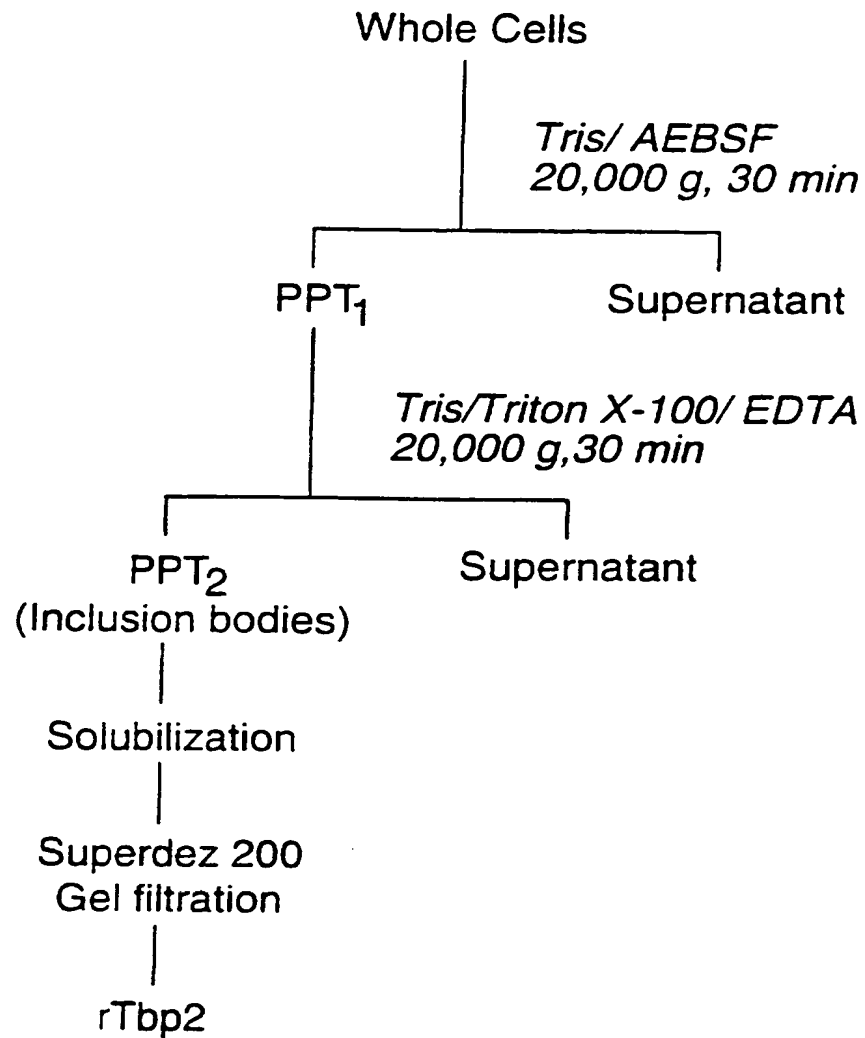
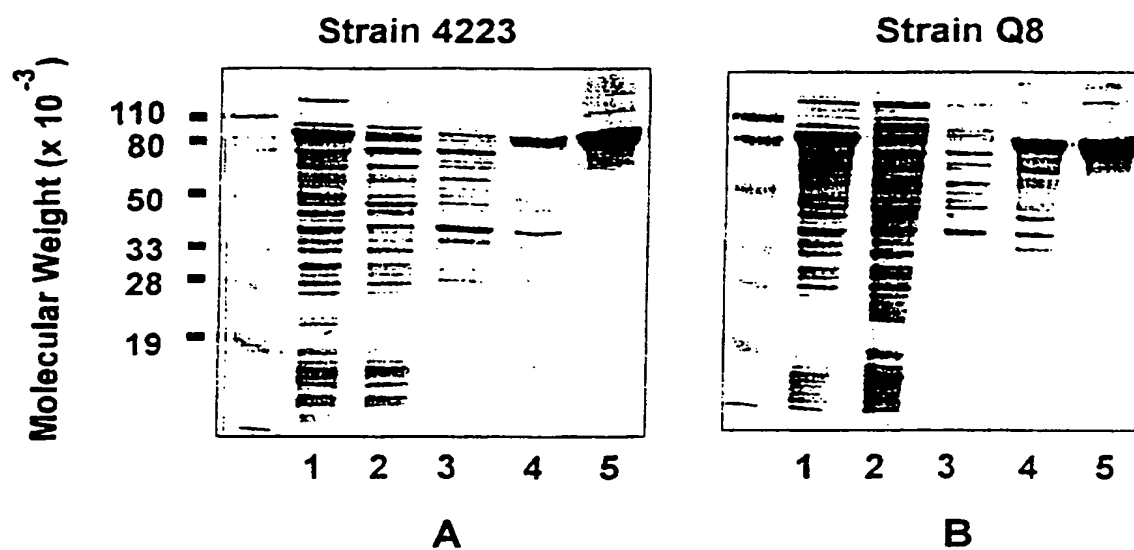
Purification of Tbp2 from *E. Coli*

FIG.22

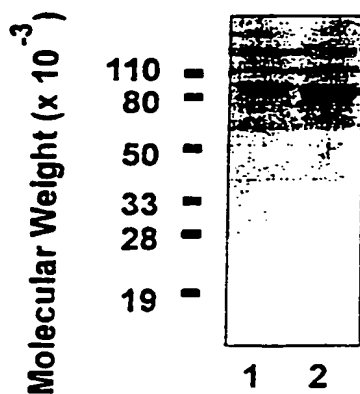
## Purification of rTbp2 from *E. coli*



1. *E. coli* Whole cells
2. Soluble proteins after 50 mM Tris extraction
3. Soluble proteins after Tris/ Triton X-100/ EDTA extraction
4. Left-over pellet (rTbp2 inclusion bodies)
5. Purified rTbp2

Fig.23

## Binding of Tbp2 to Human Transferrin



1. rTbp2 (strain 4223)

2. rTbp2 (strain Q8)

Fig.24



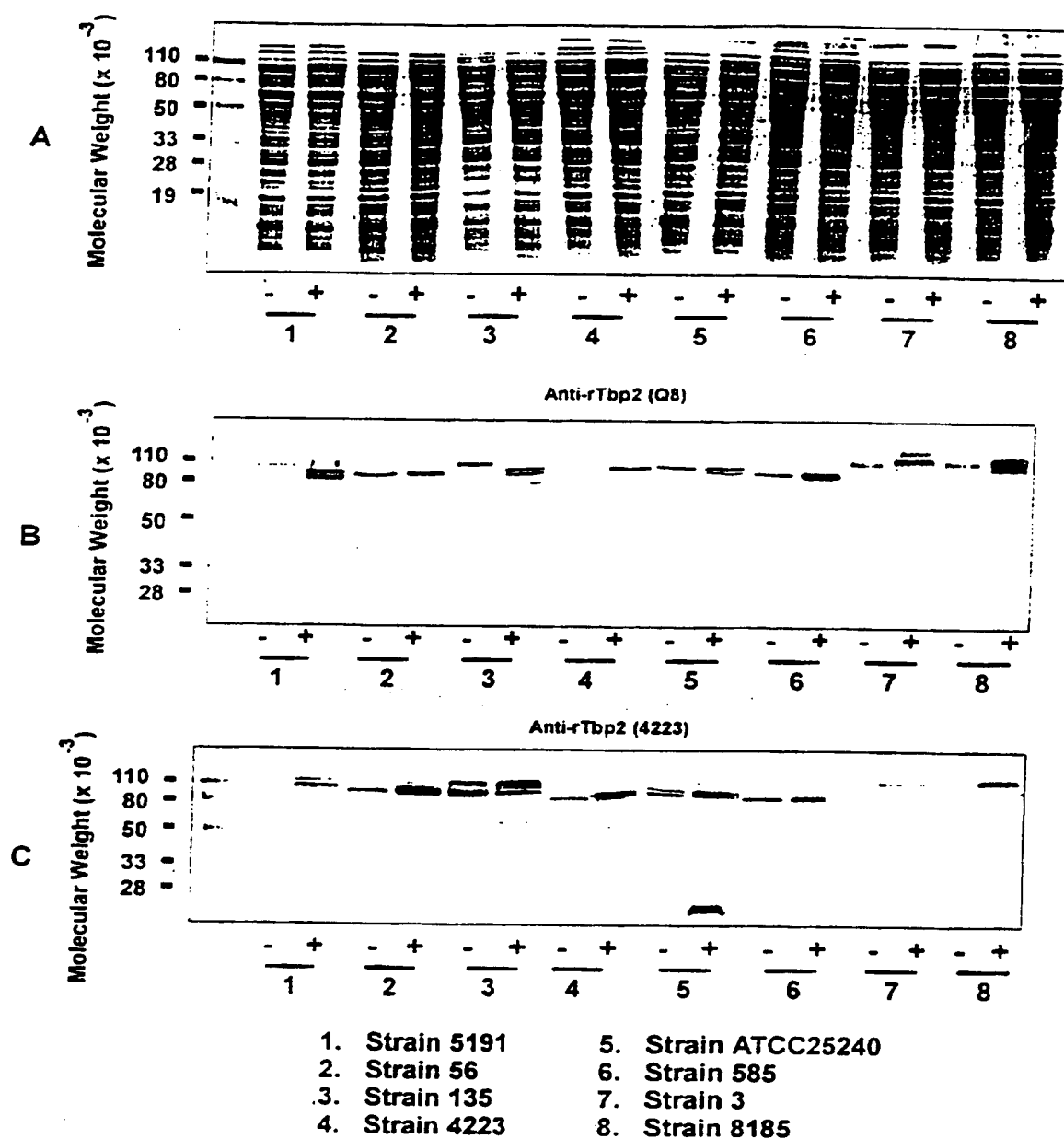


Fig.25

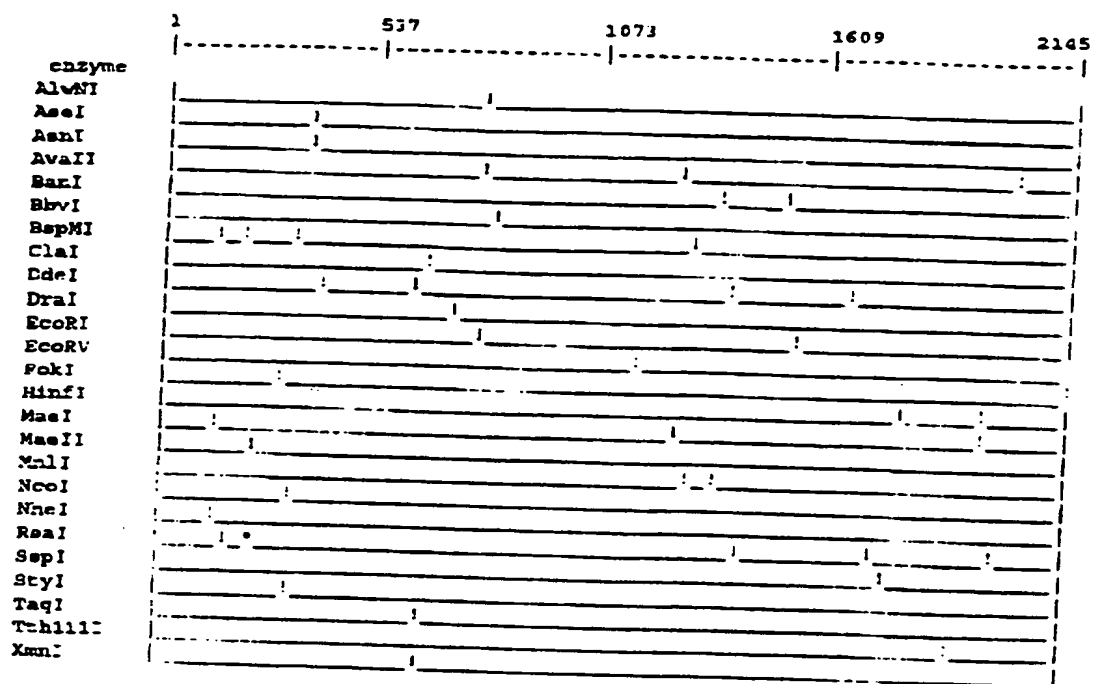
Figure 26 Restriction map of *M. catarrhalis* strain R1 *lbpB*

Figure 27 Nucleotide and deduced amino acid sequence of *M. catarrhalis* R1 *tbpB*

AAATTTGCCGTATTTTGTCTATCATAAATGCATTTATCATCAATGCCCAACAAATACGCCAAATGCACAT  
 TGTGAGCATGCCAAATAGGCATTAACAGACTTTTITAGATAATACCATCAACCCATCAGAGGATTATTTT  
 27 54  
 ATG AAA CAC ATT CCT TTA ACC ACA CTG TGT GTG GCA ATC TCT GCC GTC TTA TTA  
 MET Lys His Ile Pro Leu Thr Thr Leu Cys Val Ala Ile Ser Ala Val Leu Leu  
 81 108  
 ACC GCT TGT GGT GGC AGT GGT GGT TCA AAT CCA CCT GCT CCT ACG CCC ATT CCA  
 Thr Ala Cys Gly Gly Ser Gly Gly Ser Asn Pro Pro Ala Pro Thr Pro Ile Pro  
 135 162  
 AAT GCT AGC GGT TCA GGT AAT ACT GGC AAC ACT GGT AAT GCT GGC GGT ACT GAT  
 Asn Ala Ser Gly Ser Gly Asn Thr Gly Asn Thr Gly Asn Ala Gly Gly Thr Asp  
 189 216  
 AAT ACA GCC AAT GCA GGT AAT ACA GGC GGT ACA ASC TCT GGT ACA GGC AGT GCC  
 Asn Thr Ala Asn Ala Gly Asn Thr Gly Gly Thr Ser Ser Gly Thr Gly Ser Ala  
 243 270  
 AGC ACG TCA GAA CCA AAA TAT CAA GAT GTG CCA ACA ACG CCC AAT AAC AAA GAA  
 Ser Thr Ser Glu Pro Lys Tyr Gln Asp Val Pro Thr Thr Pro Asn Asn Lys Glu  
 297 324  
 CAA GTT TCA TCC ATT CAA GAA CCT GCC ATG GGT TAT GGC ATG GCT TTG AGT AAA  
 Gln Val Ser Ser Ile Gln Glu Pro Ala MET Gly Tyr Gly MET Ala Leu Ser Lys  
 351 378  
 ATT AAT CTA TAC GAC CAA CAA GAC ACG CCA TTA GAT GCA AAA AAT ATC ATT ACC  
 Ile Asn Leu Tyr Asp Gln Gln Asp Thr Pro Leu Asp Ala Lys Asn Ile Ile Thr  
 405 432  
 TTA GAC GGT AAA AAA CAA GTT GCT GAC AAT CAA AAA TCA CCA TTG CCA TTT TCG  
 Leu Asp Gly Lys Lys Gln Val Ala Asp Asn Gln Lys Ser Pro Leu Pro Phe Ser  
 459 486  
 TTA GAT GTA GAA AAT AAA TTG CTT GAT GGC TAT ATA GCA AAA ATG AAT GAA GCG  
 Leu Asp Val Glu Asn Lys Leu Leu Asp Gly Tyr Ile Ala Lys MET Asn Glu Ala  
 513 540  
 GAT AAA AAT GCC ATT GGT GAA AGA ATT AAG AGA GAA AAT GAA CAA AAT AAA AAA  
 Asp Lys Asn Ala Ile Gly Glu Arg Ile Lys Arg Glu Asn Glu Gln Asn Lys Lys  
 567 594  
 ATA TCC GAT GAA GAA CTT GCC AAA AAA ATC AAA GAA AAT GTG CGT AAA AGC CCT  
 Ile Ser Asp Glu Glu Leu Ala Lys Lys Ile Lys Glu Asn Val Arg Lys Ser Pro  
 621 648  
 GAG TTT CAG CAA GTA TTA TCA TCG ATA AAA GCG AAA ACT TTC CAT TCA AAT GAC  
 Glu Phe Gln Gln Val Leu Ser Ser Ile Lys Ala Lys Thr Phe His Ser Asn Asp

F1927 (cont.)

675 702  
 AAA ACA ACC AAA GCA ACC ACA CGA GAT TTA AAA TAT GTT GAT TAT GGT TAC TAC  
 Lys Thr Thr Lys Ala Thr Thr Arg Asp Leu Lys Tyr Val Asp Tyr Gly Tyr Tyr

729 756  
 TTG GTG AAT GAT GGC AAT TAT CTA ACC GTC AAA ACA GAC AAC CCA AAA CTT TGG  
 Leu Val Asn Asp Ala Asn Tyr Leu Thr Val Lys Thr Asp Asn Pro Lys Leu Trp

783 810  
 AAT TCA GGT CCT GTG GGC GGT GTG TTT TAT AAT GGC TCA ACG ACC GCC AAA GAG  
 Asn Ser Gly Pro Val Gly Gly Val Phe Tyr Asn Gly Ser Thr Thr Ala Lys Glu

837 864  
 CIG CCC ACA CAA GAT GCG GTC AAA TAT AAA GGA CAT TGG GAC TTT ATG ACC GAT  
 Leu Pro Thr Gln Asp Ala Val Lys Tyr Lys Gly His Trp Asp Phe MET Thr Asp

891 918  
 GTT GCC AAA AAA AGA AAC CGA TTT AGC GAA GTA AAA GAA ACC TAT CAA GCA GGC  
 Val Ala Lys Lys Arg Asn Arg Phe Ser Glu Val Lys Glu Thr Tyr Gln Ala Gly

945 972  
 TGG TGG TAT GGG SCA TCT TCA AAA GAT GAA TAC AAC CSC TTA TTA ACC AAA GCA  
 Trp Trp Tyr Gly Ala Ser Ser Lys Asp Glu Tyr Asn Arg Leu Leu Thr Lys Ala

999 1026  
 GAT GCC SCA CCT GAT AAT TAT AGC GGT GAA TAT GGT CAT AGC AGT GAA TTT ACT  
 Asp Ala Ala Pro Asp Asn Tyr Ser Gly Glu Tyr Gly His Ser Ser Glu Phe Thr

1053 1080  
 GTT AAT TTT AAG GAA AAA AAA TTA ACA GGT GAG CTG TTT AGT AAC CTA CAA GAC  
 Val Asn Phe Lys Glu Lys Lys Leu Thr Gly Glu Leu Phe Ser Asn Leu Gln Asp

1107 1134  
 AGC CAT AAA CAA AAA GTA ACC AAA ACA AAA CGC TAT GAT ATT AAG GCT GAT ATC  
 Ser His Lys Gln Lys Val Thr Lys Thr Lys Arg Tyr Asp Ile Lys Ala Asp Ile

1161 1188  
 CAC GGC AAC CGC TTC CGT GGC AGT GCC ACC GCA AGC GAT AAG GCA GAA GAC AGC  
 His Gly Asn Arg Phe Arg Gly Ser Ala Thr Ala Ser Asp Lys Ala Glu Asp Ser

1215 1242  
 AAA AGC AAA CAC CCC TTT ACC ACC GAT GCC AAA GAT AAG CTA GAA GGT GGT TTT  
 Lys Ser Lys His Pro Phe Thr Ser Asp Ala Lys Asp Lys Leu Glu Gly Gly Phe

1269 1296  
 TAT GGA CCA AAA GGC GAG GAG CTG GCA GGT AAA TTC TTA ACC GAT GAT AAC AAA  
 Tyr Gly Pro Lys Gly Glu Glu Leu Ala Gly Lys Phe Leu Thr Asp Asp Asn Lys

1323 1350  
 CTC TTT GGT GTC TTT GGT GCC AAA CAA GAG GGT AAT GTA GAA AAA ACC GAA GCC  
 Leu Phe Gly Val Phe Gly Ala Lys Gln Glu Gly Asn Val Glu Lys Thr Glu Ala

1377 1404  
 ATC TTA GAT GGT TAT GCA CTT GGS ACA TTT AAT AAA CCT GGT ACG ACC AAT CCC  
 Ile Leu Asp Ala Tyr Ala Leu Gly Thr Phe Asn Lys Pro Gly Thr Thr Asn Pro

1431 1458  
 GCC TTT ACC GCT AAC AGC AAA AAA GAA CTG GAT AAC TTT GGC AAT GCC AAA AAG  
 Ala Phe Thr Ala Asn Ser Lys Lys Glu Leu Asp Asn Phe Gly Asn Ala Lys Lys

1485 1512  
 TTG GTC TTG GGT TCT ACC GTC ATT GAT TTG GTG CCT ACT GAT GUC ACC AAA GAT  
 Leu Val Leu Gly Ser Thr Val Ile Asp Leu Val Pro Thr Asp Ala Thr Lys Asp

1539 1566  
 GTC AAT GAA TTC AAA GAA AAG CCA AAG TCT GCC ACA AAC AAA GCG GGC GAA ACT  
 Val Asn Glu Phe Lys Glu Lys Pro Lys Ser Ala Thr Asn Lys Ala Gly Glu Thr

1593 1620  
 CTG ATG GTG AAT GAT GAA GTT AGC GTC AAA ACC TAT GGC AAA AAC TTT GAA TAC  
 Leu MET Val Asn Asp Glu Val Ser Val Lys Thr Tyr Gly Lys Asn Phe Glu Tyr

1647 1674  
 CTA AAA TTT GGT GAG CTT AGT GTC GGT GGT AGC CAT AGC GTC TTT TTA CAA GGC  
 Leu Lys Phe Gly Glu Leu Ser Val Gly Gly Ser His Ser Val Phe Leu Gln Gly

1701 1728  
 GAA CGC ACC GGT ACC ACA GGC GAG AAA GCC GTA CCA ACC ACA GGC AAA GCC AAA  
 Glu Arg Thr Ala Thr Thr Gly Glu Lys Ala Val Pro Thr Thr Gly Lys Ala Lys

1755 1782  
 TAT TTG GGG AAC TGG GTA GGA TAT ATC ACA GGA GCG GAC TCA TCA AAA GGC TCT  
 Tyr Leu Gly Asn Trp Val Gly Tyr Ile Thr Gly Ala Asp Ser Ser Lys Gly Ser

1809 1836  
 ACC GAT GGC AAA GGC TTT ACC GAT GCC AAA GAT ATT GCT GAT TTT GAC ATT GAC  
 Thr Asp Gly Lys Gly Phe Thr Asp Ala Lys Asp Ile Ala Asp Phe Asp Ile Asp

1863 1890  
 TTT GAG AAA AAA TCA GTT AAT GGC AAA CTG ACC ACC AAA GAC CGC CAA GAC CCT  
 Phe Glu Lys Lys Ser Val Asn Gly Lys Leu Thr Thr Lys Asp Arg Gln Asp Pro

1917 1944  
 GTC TTT AAC ATC ACA GGT GAA ATC GCA GGC AAT GGC TGG ACA GGT AAA GCC AGC  
 Val Phe Asn Ile Thr Gly Glu Ile Ala Gly Asn Gly Trp Thr Gly Lys Ala Ser

1971 1998  
 ACC GCC GAA GCG AAC GCA GGG GGC TAT AAG ATA GAT TCT AGC AGT ACA GGC AAA  
 Thr Ala Glu Ala Asn Ala Gly Gly Tyr Lys Ile Asp Ser Ser Ser Thr Gly Lys

2025 2052  
 TCC ATC GTC ATC AAA GAT GCC GTG GTT ACA GGT GGC TTT TAT GGT CCA AAT GCA  
 Ser Ile Val Ile Lys Asp Ala Val Val Thr Gly Gly Phe Tyr Gly Pro Asn Ala

89/90

Fig 27 (cont)

2079

ACC GAG ATG GGT GGG TCA TTT ACA CAC AAC AGC GGT AAT GAT GGT AAA CTC TCT  
Thr Glu MET Gly Gly Ser Phe Thr His Asn Ser Gly Asn Asp Gly Lys Val Ser

2106

2133

GTG CTC TTT GGC ACA AAA AAA CAA GAA GTT AAG AAG TGA  
Val Val Phe Gly Thr Lys Lys Gln Glu Val Lys Lys \*



# INTERNATIONAL SEARCH REPORT

In national Application No  
CA 97/00163

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/12 C07K14/22 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 97 13785 A (CONNAUGHT LAB ; YANG YAN PING (CA); MYERS LISA E (CA); HARKNESS ROB) 17 April 1997 see the whole document	1-25
Y	WO 90 12591 A (UNIV TECHNOLOGIES INTERNATIONAL ; SCHRYVERS ANTHONY BERNARD (CA)) 1 November 1990 see claims 1-26	1-25
	- / - -	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*&\* document member of the same patent family

Date of the actual completion of the international search

17 July 1997

Date of mailing of the international search report

30 JULY 1997 (30.07.97)

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

Nauche, S



## INTERNATIONAL SEARCH REPORT

International Application No.

/CA 97/00163

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MICROBIAL PATHOGENESIS, vol. 15, 1993, pages 433-445, XP000612196 RAONG-HUA YU ET AL: "THE INTERACTION BETWEEN HUMAN TRANSFERRIN AND TRANSFERRIN BINDING PROTEIN 2 FROM MORAXELLA (BRANHAMELLA) CATARRHALIS DIFFERS FROM THAT OF OTHER HUMAN PATHOGENS" see the whole document ---	1-25
A	WO 95 33049 A (PASTEUR MERIEUX SERUMS VACC ;TRANSGENE SA (FR); MILLET MARIE JOSE) 7 December 1995 see the whole document ---	1-25
A	WO 93 08283 A (UNIV SASKATCHEWAN) 29 April 1993 see the whole document -----	1-25

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA 97/ 00163

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 23  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claim 23 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

/CA 97/00163

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9713785 A	17-04-97	AU 7208296 A	30-04-97
WO 9012591 A	01-11-90	US 5292869 A	08-03-94
		AU 649950 B	09-06-94
		AU 5526190 A	16-11-90
		EP 0528787 A	03-03-93
		JP 4506794 T	26-11-92
		NZ 247967 A	24-02-95
		US 5141743 A	25-08-92
WO 9533049 A	07-12-95	FR 2720408 A	01-12-95
		AU 2675795 A	21-12-95
		CA 2167936 A	07-12-95
		EP 0720653 A	10-07-96
		FI 960428 A	28-03-96
		HU 75992 A	28-05-97
		JP 9501059 T	04-02-97
		NO 960332 A	21-03-96
WO 9308283 A	29-04-93	US 5417971 A	23-05-95
		AU 2751392 A	21-05-93
		CA 2121364 A	29-04-93
		EP 0610260 A	17-08-94
		US 5521072 A	28-05-96

**THIS PAGE BLANK (USPTO,**